Patched Genes and Uses Related Thereto

Related Applications

This application is a continuation-in-part of U.S.S.N. 08/656,055, which is a continuation-in-part of U.S.S.N. 08/540,406, which is a continuation-in-part of U.S.S.N. 08/31~~7~~9,745 (now abandoned). The specifications of each of these prior applications are incorporated herein by reference.

RS 5/7/02

Background of the Invention

Segment polarity genes were originally discovered as mutations in flies that change the pattern of body segment structures. Mutations in these genes cause animals to develop changed patterns on the surfaces of body segments; the changes affecting the pattern along the head to tail axis. Among the genes in this class are *hedgehog,* which encodes a secreted protein (HH), and *patched,* which encodes a protein structurally similar to transporter proteins, having twelve transmembrane domains (*ptc*), with two conserved glycosylation signals.

The *hedgehog* gene of flies has at least three vertebrate relatives- *Sonic hedgehog (Shh); Indian hedgehog (Ihh), and Desert hedgehog (Dhh). Shh* is expressed in a group of cells, at the posterior of each developing limb bud, that have an important role in signaling polarity to the developing limb. The *Shh* protein product, SHH, is a critical trigger of posterior limb development, and is also involved in polarizing the neural tube and somites along the dorsal ventral axis. Based on genetic experiments in flies, *patched* and *hedgehog* have antagonistic effects in development. The *patched* gene product, *ptc*, is widely expressed in fetal and adult tissues, and plays an important role in regulation of development. *Ptc* downregulates transcription of itself, members of the transforming growth factor and *Wnt* gene families, and possibly other genes. Among other activities, HH upregulates expression of *patched* and other genes that are negatively regulated by *patched.*

It is of interest that many genes involved in the regulation of growth and control of cellular signaling are also involved in oncogenesis. Such genes may be oncogenes, which are typically upregulated in tumor cells, or tumor suppressor genes, which are down-regulated or absent in tumor cells. Malignancies may arise when a tumor suppressor is lost and/or an oncogene is inappropriately activated. Familial predisposition to cancer may occur when there is a mutation, such as loss of an allele encoding a suppressor gene, present in the germline DNA of an individual.

The most common form of cancer in the United States is basal cell carcinoma of the skin. While sporadic cases are very common, there are also familial syndromes, such as the basal cell nevus syndrome (BCNS). The familial syndrome has many features indicative of abnormal embryonic development, indicating that the mutated gene also plays an important role in development of the embryo. A loss of heterozygosity of chromosome 9q alleles in both familial and sporadic carcinomas suggests that a tumor suppressor gene is present in this region. The high incidence of skin cancer makes the identification of this putative tumor suppressor gene of great interest for diagnosis, therapy, and drug screening.

Relevant Literature

Descriptions of *patched,* by itself or its role with *hedgehog* may be found in Hooper and Scott (1989) Cell **59**-.751-765; and Nakano *et al.* (1989) Nature **341** -.508-513. Both of these references also describe the sequence for *Drosophila patched.* Discussions of the role of *hedgehog* include Riddle *et al.* (1993) Cell 75-.1401-1416-, Echelard *et al.* (1993) Cell **75**:1417-1430- Krauss *et al.* (1993) Cell **75**:1431-1444 (1993); Tabata and Kornberg (1994) 76:89-102; Heemskerk and DiNardo (1994) Cell **76**:449-460; and Roelink *et al.* (1994) Cell **76**:-761-775.

Mapping of deleted regions on chromosome 9 in skin cancers is described in Habuchi *et al.* (1995) Oncogene **11**: 1 671-1674, Quinn *et al.* (1 994) Genes Chromosome Cancer **11**:222-225; Quinn *et al.* (1994) J. Invest. Dermatol. 102:300-303; and Wicking *et al.* (1994) Genomics **22**:505-51 1.

Gorlin (1987) Medicine **66**:98-113 reviews nevoid basal cell carcinoma syndrome. The syndrome shows autosomal dominant inheritance with probably complete penetrance. About 60% of the cases represent new mutations. Developmental abnormalities found with this syndrome include rib and craniofacial abnormalities, polydactyly, syndactyly and spina bifida. Tumors found with the syndrome include basal cell carcinomas, fibromas of the ovaries and heart, cysts of the skin, jaws and mesentery, meningiomas and medulloblastomas.

Summary of the Invention

Isolated nucleotide compositions and sequences are provided for *patched (ptc)* genes, including mammalian, e.g. human and mouse, and invertebrate homologs. Decreased expression of *ptc* is associated with the occurrence of human cancers, particularly basal cell carcinomas and other tumors of epithelial tissues such as the skin. The cancers may be familial, having as a component of risk a germline mutation in the gene, or may be sporadic. *Ptc*, and its antagonist *hedgehog,* are useful in creating transgenic animal models for these human cancers. The *ptc* nucleic acid compositions find use in identifying homologous or

related genes; in producing compositions that modulate the expression or function of its encoded protein, *ptc*; for gene therapy; mapping functional regions of the protein- and in studying associated physiological pathways. In addition, modulation of the gene activity *in vivo* is used for prophylactic and therapeutic purposes, such as treatment of cancer, identification of cell type based on expression, and the like. *Ptc*, anti-*ptc* antibodies and *ptc* nucleic acid sequences are useful as diagnostics for a genetic predisposition to cancer or developmental abnormality syndromes, and to identify specific cancers having mutations in this gene.

Brief Description of the Drawings

Fig. 1 is a graph having a restriction map of about 10 kbp of the 5' region upstream from the initiation codon of *Drosophila patched* gene and bar graphs of constructs of truncated portions of the 5' region joined to fl-galactosidase, where the constructs are introduced into fly cell lines for the production of embryos. The expression of fl-gal in the embryos is indicated in the right-hand table during early and late development of the embryo. The greater the number of +'s, the more intense the staining.

Fig. 2 shows a summary of mutations found in the human *patched* gene locus that are associated with basal cell nevus syndrome. Mutation (1) is found in sporadic basal cell carcinoma, and is a C to T transition in exon 3 at nucleotide 523 of the coding sequence, changing Leu 175 to Phe in the first extracellular loop. Mutations 2-4 are found in hereditary basal carcinoma nevus syndrome. (2) is an insertion of 9 bp at nucleotide 2445, resulting in the insertion of an additional 3 amino acids after amino acid 815. (3) is a deletion of 11 bp, which removes nt 2442-2452 from the coding sequence. The resulting frameshift truncates the open reading frame after amino acid 813, 'just after the seventh transmembrane domain. (4) is a G to C alteration that changes two conserved nucleotides of the 3' splice site adjacent to exon 10, creating a non-functional splice site that truncates the protein after amino acid 449, in the second transmembrane region.

Fig. 3 (panels A-B) illustrates the generation of *ptc* mutations. **(A)** The *ptc* mutant allele was generated by homologous recombination between the KO1 targeting vector and *ptc*. External probe A detected a 3' EcoRV polymorphism on blots and probe B detected a 5' SacI polymorphism. Exons are numbered. **(B)** Transmission of the ptc^{KO1} allele through the germline was confirmed by Southern blot (upper panel) and a PCR genotyping assay (lower panel). PCR primers are indicated as arrows in **A**. Because the homozygous mutant embryos were being resorbed, there was much less yolk sac DNA in the -/- lanes.

Fig. 4 (panels A-G) illustrate the germ layer-specific derepression of Hh target genes in $ptc^{-/-}$ embryos. **(A, B)** Lateral views of E8.25 wild-type **(A)** and $ptc^{-/-}$ **(B)** embryos. The

headfolds are overgrown in the mutant (white arrows) and the heart is not properly formed (red arrows). **(C)** Lateral views of E8.75 *ptc*$^{+/-}$ (left) and *ptc*$^{-/-}$ (right) embryos stained with X-gal *(28)* **(D, E, F, G)** Transverse sections through E8.75 *ptc*$^{+/-}$ **(D, F)** and *ptc*$^{-/-}$ **(E, G)** embryos stained with X-gal **(D, E)** or hybridized with a digoxigenin labeled *Gli* probe *(29)* **(F, G)**. Both *lacZ* and *Gli* were derepressed in the ectoderm and mesoderm but not in the endoderm (arrows). In **A** and **B**, anterior is to the left and dorsal is up. In **C**, anterior is up and dorsal is to the right. In **D** to **G**, dorsal is up.

Fig. 5 (panels A-L) illustrate ventralization of the neural tube in *ptc*$^{-/-}$ embryos. **(A)** Lateral view of E8.5 wild-type (left) and *ptc*$^{-/-}$ (right) embryos hybridized with a *HNF3b* probe. Expression is expanded dorsally in the mutant. **(B, C)** Transverse sections through the hindbrain of E8.5 wild-type **(B)** and *ptc*$^{-/-}$ **(C)** embryos hybridized with ^{35}S-labeled *Shh* probe *(8)*. *Shh* is expressed in the floor plate (fp) and notochord (nc) of the wild-type embryo, and is greatly expanded in the *ptc* mutant . g = gut **(D, E)** Hematoxylin and eosin stained transverse sections through, the hindbrain of wild-type **(D)** and *ptc*$^{-/-}$ **(E)** E8.5 embryos. Bottle-shaped cells with basal nuclei are indicated by arrows. **(F, G)** Transverse sections through E8.5 *ptc*$^{+/-}$ **(F)** and *ptc*$^{-/-}$ **(G)** embryos hybridized with *Pax6* probe show loss of expression from the *ptc* mutant. **(H)** Dorsal view of E8.25-E8.5 embryos hybridized with *Pax3* probe. Because of the kinking in the neural tube, the *ptc*$^{-/-}$ embryo is curled on itself. Weak *Pax3* expression is seen in the posterior dorsal neural tube of the *ptc*$^{-/-}$ embryo (bottom, arrow). **(I, J)** Transverse sections through E8.5 wild-type **(I)** and *ptc*$^{-/-}$ **(J)** embryos hybridized with *Pax3* probe. *Pax3* is expressed in the dorsal neural tube (nt) and dermamyotome (dm) in the wild-type, but is only present in a small dorsal domain of the mutant neural tube. s = somite **(K, L)** Lateral views of E9 wild type **(K)** and E8.5 *ptc*$^{-/-}$ **(L)** embryos hybridized with *erb-b3* probe. Staining is seen in migrating neural crest in the head and somites of wild type but not mutant embryos (red arrows). Weak staining in the head, heart and gut (black arrows) is background or non-neural crest related. **(M)** Lateral view of wild type (top) and *ptc*$^{-/-}$ (bottom) embryos hybridized with *Nkx2.1* probe. The body of the mutant is twisted. *Nkx2.1* expression is limited to the anterior, but is expanded dorsally in the mutant. **(N)** Lateral view of E8.5 *ptc*$^{+/-}$ (left) and *ptc*$^{-/-}$ (right) embryos hybridized with *hoxb1* probe. Loss of expression in rhombomere four is indicated by the asterisks. In all transverse sections, dorsal is up. In **A**, **K**, **L** and **N**, anterior is up and dorsal is to the right. In **H** and **M**, anterior is to the left.

RS 6/16/02

Fig. 6 (panels A-F) depict s keletal abnormalities and medulloblastomas in *ptc*$^{+/-}$ mice **(A)** Alcian blue and Alizarin red stained hindlimb from a *ptc*$^{+/-}$ mouse *(30)*. The preaxial digit is duplicated (arrows). **(B, C)** Dorsal views of brains from wild-type **(B)** and *ptc*$^{+/-}$ **(C)** mice. Anterior is up. In the posterior wild-type brain, the colliculi (col) are present as distinct bumps between the cortex (cor) and cerebellum (ce). In the *ptc*$^{+/-}$ mouse, a massive medulloblastoma (mb, outlined in red) grew over the colliculi and normal

cerebellum, which can no longer be seen. The olfactory bulbs were removed. **(D, E)** Hematoxylin and eosin stained section through human **(D)** and mouse **(E)** medulloblastomas. The tumor cells are small with dark, carrot-shaped nuclei (arrows) and form nodules with no apparent orientation. **(F)** Synaptophysin immunoreactivity in a mouse medulloblastoma *(26)*. Synaptophysin staining (brown) is seen in some processes (arrows). Nuclei are purple.

Fig. 7 (panels A-G) illustrate derepression of *ptc* and *Gli* expression in medulloblastomas from *ptc*$^{+/-}$ mice. **(A to C)** Semi-adjacent sections through a tumor in the cerebellum of a *ptc*$^{+/-}$ mouse hybridized with ^{35}S labeled probes to *ptc* **(A)**, *Gli* **(B)** and *Shh* **(C)**. *ptc* and *Gli* transcripts are abundant in the tumors (asterisks) compared to nearby cerebellar tissue (arrows). No *Shh* was detected in the tumor. **(D)** *ptc*$^{+/-}$ cerebellum (ce) and tumor (mb) stained with X-gal *(28)*. Anterior is to the left. Derepression of *ptc* expression in the medulloblastoma is reflected in the high level of X-gal staining. **(E)** Surface staining in (arrows) regions of *ptc*$^{+/-}$ cerebellum contrast with absence of b-galactosidase activity in most folia (asterisk). **(F)** Sagittal section through cerebellum in **E**. X-gal staining nuclei (arrow) accumulated superficial to the molecular layer (ml), where stained nuclei are not normally seen. In unaffected regions of the cerebellum, X-gal staining was seen in scattered cells of the molecular layer (ml), strongly in the Purkinje cell layer (pcl) and weakly in the granule cell layer (gl). **(G)** *ptc* expression was examined in total RNA (15 mg) from wild-type (WT) and *ptc*$^{+/-}$ (+/-) cerebellums using a probe (M2-2) *(6)* that detects exons downstream of the *lacZ* and *neo* insertions. Actin mRNA was used as an RNA loading control. The *ptc*$^{+/-}$ mice had~50% decrease in *ptc* transcripts.

Database References for Nucleotide and Amino Acid Sequences

The sequence for the *D. melanogaster patched* gene has the Genbank accession number M28418. The sequence for the mouse *patched* gene has the Genbank accession number lt30589-V46155. The sequence for the human *patched* gene has the Genbank accession number U59464.

Detailed Description of the Invention

Vertebrate and invertebrate *patched (ptc)* gene compositions and methods for their isolation are provided. Of particular interest are mammalian *ptc* genes, such as the human and mouse homologs described in the appended examples. The *ptc* gene, in mammals, is a tumor suppressor and developmental regulator. Certain human cancers, e.g. basal cell carcinoma, transitional cell carcinoma of the bladder, meningiomas, medulloblastomas, etc., can be characterized by *ptc* loss-of-function, such as that resulting from oncogenic mutations at the *ptc* locus, or other loss-of-function mutations which decrease *ptc* activity in the cell. As

described below, we have observed somatic mutations in the *ptc* gene in a variety of sporadic cancers. For instance, the basal cell nevus syndrome (BCNS), an inherited disorder, is associated with germline mutations in *ptc*. Some patients with basal cell nevus syndrome (BCNS) have germ line mutations in *ptc*, and are at increased risk for developmental defects such as spina bifida and craniofacial abnormalities, basal cell carcinoma (BCC) of the skin, and brain tumors. Mutations to *ptc* genes are also observed to occur in sporadic BCCs, which generally have both copies of *ptc* inactivated.

The term "loss-of-function" is art recognized and, with respect to a *patched* gene or gene product refers to mutations in a *patched* gene which ultimately decrease or otherwise inhibit the ability of a cell to transduce *patched*-mediated signals, e.g., the cells may lose responsiveness to *hedgehog* induction. For example, a loss-of-function mutation to a *patched* gene may be a point mutation, deletion or insertion of sequences in the coding sequence, intron sequence or 5' or 3' flanking sequences of the gene so as to, for example, (i) alter (e.g., decrease) the level *patched* expression, (ii) alter exon-splicing patterns, (iii) alter the ability of the encoded *patched* protein to interact with extracellular or intracellular proteins (such as *hedgehog*), or (iv) alter (decrease) the stability of the encoded *patched* protein.

The term "aberrant modification" is art recognized and, with respect to a *patched* gene, refers to a non-wildtype mutation or other alteration to the gene, e.g., which results in full or partial loss-of-function of the *patched* protein or expression of the *patched* gene.

RS 5/6/02

Such mutations affecting *ptc* activity have also been associated with other human cancers, including carcinomas, adenocarcinomas, sarcomas and the like. Decreased *ptc* activity is also associated with inherited developmental abnormalities, e.g. rib and craniofacial abnormalities, polydactyly, syndactyly and spina bifida.

The art-recognized term "predisposing mutation", as it pertains to *patched* genes, refers to mutations to the *patched* gene which result in loss-of-function.

The term "genetic predisposition" is art recognized, and refers to a genotype of an animal which predisposes the animal to developing a certain pathological conditions with a frequency (probability) greater than the average for the overall population of that animal, taking into account, as appropriate, age, sex or other related physical or medical condition(s).

The *ptc* genes and fragments thereof, encoded protein, and anti-*ptc* antibodies are useful in the identification of individuals predisposed to development of a variety of cancers and developmental abnormalities, and in characterizing the phenotype of various tumors or other proliferative or degenerative disorders that are associated with this gene, e.g., for diagnostic and/or prognostic benefit. The characterization is useful for prenatal screening; and in determining the phenotype of a proliferative disorder, e.g. for determining a course of treatment of the patient. Tumors may be typed or staged as to the *ptc* status, e.g. by detection

of mutated sequences, antibody detection of abnormal protein products, and functional assays for altered *ptc* activity.

The terms "developmental disorder" and "develpmental abnormality" are art recognized, and refer to abberant development of a cell, tissue or organ, e.g., in size, symmetry or functional performance, which abnormality may or may not be untowardly manifest.

The term "proliferative disorder" is art recognized and refers to a disorder affecting an animal in a manner which is marked by abberant, or otherwise unwanted, proliferation of a subset of cells of an animal. Cancers are proliferative disorders.

The encoded *ptc* protein is also useful in drug screening for compositions that mimic *ptc* activity or expression, including altered forms of *ptc* protein, particularly with respect to *ptc* function as a tumor suppressor in oncogenesis.

The human and mouse *ptc* gene sequences and isolated nucleic acid compositions are provided in the appended examples. In identifying the mouse and human *patched* genes, cross-hybridization of DNA and amplification primers were employed to move through the evolutionary tree from the known *Drosophila ptc* sequence, identifying a number of invertebrate homologs.

Sub F1 The human *patched* gene has been mapped to human chromosome band 9q22.3, and lies between the polymorphic markers D9S196 and D9S287 (a detailed map of human genome markers may be found in Dib *et al.* (1996) Nature 280:152- http://www.genethon.fr).

As will be understood by those skilled in the art, the method of the present invention can be carried out using any of a large number of assay techniques for detecting alterations in *ptc* genes and/or *ptc* protein function. For instance, individuals are screened by analyzing their DNA or RNA for the presence of a predisposing oncogenic or developmental mutation, as compared to a normal sequence. An exemplary "normal" sequence of *patched* is provided in SEQ ID NO:19 (human). Specific mutations of interest include any mutation that leads to oncogenesis or developmental abnormalities, including insertions, substitutions and deletions in the coding region sequence, in the introns (e.g., that affect splicing), in the transcriptional regulatory sequences (such as promoter or enhancer sequences) that affect the activity and expression of the protein.

In general, the subject method can be characterized as including a step of detecting, in a sample of cells from a patient, the presence or absence of *ptc* expression (at the protein or mRNA transcript level), mutations to the *ptc* gene (coding or non-coding sequence) and/or the functional activity of *ptc* in the sample of cells (such as induction of Gli or the like). Moreover, the subject method can be used to assess the phenotype of cells which are known

to be transformed, the phenotype results being useful in planning a particular therapeutic regimen.

To illustrate, nucleic acid samples are obtained from a patient having, or suspected as being at risk for developing, a tumor or developmental abnormality which may be associated with *ptc*. The nucleic acid is analyzed for the presence of a predisposing mutation in the *ptc* gene. The presence of a mutated *ptc* sequence that affects the level of expression of the gene, stability of the gene product, and/or signal transduction activity of *ptc* confers an increased susceptibility to a proliferative or developmental disorder. Thus, the level of expression of *ptc* can be used predictively to evaluate whether a sample of cells contains cells which are, or are predisposed towards becoming, transformed.

Diagnostic/prognostic screening of tissue/cell samples for tumors or developmental abnormalities may also be based on the functional or antigenic characteristics of the protein. Immunoassays designed to detect the normal or abnormal *ptc* protein may be used in screening. Where many diverse mutations lead to a particular disease phenotype, functional protein assays have proven to be effective screening tools. Such assays may be based on detecting changes in the transcriptional regulation mediated by *ptc*, or may directly detect *ptc* activities such as *hedgehog* binding, transporter activity or the like, or may involve antibody localization of *patched* in cells.

Inheritance of BCNS is autosomal dominant, although many cases are the result of new mutations. Diagnosis of BCNS is performed by protein, DNA sequence or hybridization analysis of any convenient sample from a patient, e.g. biopsy material, blood sample, scrapings from cheek, etc. A typical patient genotype will have a predisposing mutation on at least one chromosome. In tumors and at least sometimes developmentally affected tissues, loss of heterozygosity at the *ptc* locus leads to aberrant cell and tissue behavior. When the normal copy of *ptc* is lost, leaving only the reduced function mutant copy, abnormal cell growth and reduced cell layer adhesion is the result. Examples of specific *ptc* mutations in BCNS patients are a 9 bp insertion at nt 2445 of the coding sequence- and an 11 bp deletion of nt 2441 to 2452 of the coding sequence. These result in insertions or deletions in the region of the seventh transmembrane domain.

Prenatal diagnosis of BCNS may be performed, particularly where there is a family history of the disease, e.g. an affected parent or sibling. It is desirable, although not required, in such cases to determine the specific predisposing mutation present in affected family members. A sample of fetal DNA, such as an amniocentesis sample, fetal nucleated or white blood cells isolated from maternal blood, chorionic villus sample, etc. is analyzed for the presence of the predisposing mutation. Alternatively, a protein based assay, e.g. functional assay or immunoassay, is performed on fetal cells known to express *ptc*.

Sporadic tumors associated with loss of *ptc* function include a number of carcinomas and other transformed cells known to have deletions in the region of chromosome 9q22, e.g. basal cell carcinomas, transitional bladder cell carcinoma, meningiomas, medullomas, fibromas of the heart and ovary, and carcinomas of the lung, ovary, kidney and esophagus. Characterization of sporadic tumors will generally require analysis of tumor cell DNA, conveniently with a biopsy sample. A wide range of mutations are found in sporadic cases, up to and including deletion of the entire long arm of chromosome 9. Oncogenic mutations may delete one or more exons, e.g. 8 and 9, may affect the amino acid sequence such as of the extracellular loops or transmembrane domains, may cause truncation of the protein by introducing a frameshift or stop codon, etc. Specific examples of oncogenic mutations include a C to T transition at nt 523 and deletions encompassing exon 9. C to T transitions are characteristic of ultraviolet mutagenesis, as expected with cases of skin cancer.

Biochemical studies may be performed to determine whether a candidate sequence variation in the *ptc* coding region or control regions is oncogenic. For example, a change in the promoter or enhancer sequence that downregulates expression of *patched* may result in predisposition to cancer. Expression levels of a candidate variant allele are compared to expression levels of the normal allele by various methods known in the art. Methods for determining promoter or enhancer strength include quantitation of the expressed mRNA or *ptc* protein; insertion of the variant control element into a vector with a reporter gene such as β-galactosidase, chloramphenical acetyltransferase, etc. that provides for convenient quantitation- and the like. Nuclear run-off assays are another ~~anpther~~ convenient means for measuring promoter/enhancer activity. The activity of the encoded *ptc* protein may be determined by comparison with the wild-type protein, e.g. by detection of transcriptional regulation of TGF or *Wnt* family genes, Gli genes, *ptc* itself, or reporter gene fusions involving transcriptional regulatory sequences of these target genes.

RS 5/6/02

The term "*patched*-dependent gene", or "a gene which is regulated in a *patched*-dependent manner", refers to genes, such as Gli or *patched*, etc, whose level of expression is regulated at least in part by the presence of a *patched* protein in the cell, e.g., can be controlled by *patched*-dependent intracellular signals.

A human *patched* gene (SEQ ID NO:18) has a 4.5 kb open reading frame encoding a protein of 1447 amino acids. Including coding and noncoding sequences, it is about 89% identical at the nucleotide level to the mouse *patched* gene (SEQ ID NO:9). A mouse *patched* gene (SEQ ID NO:9) encodes a protein (SEO ID NO:10) that has about 38% identical amino acids to *Drosophila ptc* (SEQ ID NO:6), over about 1,200 amino acids. The butterfly homolog (SEQ ID NO:4) is 1,300 amino acids long and overall has a 50% amino acid identity to fly *ptc* (SEQ ID NO:6). A 267 bp exon from the beetle patched gene encodes an

89 amino acid protein fragment, which was found to be 44% and 51% identical to the corresponding regions of fly and butterfly *ptc* respectively.

The DNA sequence encoding *ptc* may be cDNA, RNA, genomic DNA or synthetic, an includes fragments of the full-length coding sequence. The term "*patched* gene" shall be intended to mean the open reading frame encoding specific *ptc* polypeptides, as well as, as appropriate, adjacent intronic sequences and 5' and 3' non-coding nucleotide sequences involved in the regulation of expression, up to about 1 kb beyond the coding region, in either direction. The gene may be introduced into an appropriate vector for extrachromosomal maintenance or for integration into the host.

The term "cDNA" as used herein is intended to include all nucleic acids that share the arrangement of sequence elements found in native mature mRNA species, where sequence elements are exons, 3' and 5' non-coding regions. Normally mRNA species have contiguous exons, with the intervening introns deleted, to create a continuous open reading frame encoding *ptc*.

The genomic *ptc* sequence has a non-contiguous open reading frame, where introns interrupt the coding regions. A genomic sequence of interest comprises the nucleic acid present between the initiation codon and the stop codon, as defined in the listed sequences, including all of the introns that are normally present in a native chromosome. It may further include the 3' and 5' untranslated regions found in the mature mRNA. It may further include specific transcriptional and translational regulatory sequences, such as promoters, enhancers, etc., including about 1 kb of flanking genomic DNA at either the 5' or 3' end of the coding region. The genomic DNA may be isolated as a fragment of 50 kbp or smaller; and substantially free of flanking chromosomal sequence.

The nucleic acid compositions of the subject invention encode all or a part of the subject polypeptides. Fragments may be obtained of the DNA sequence by chemically synthesizing oligonucleotides in accordance with conventional methods, by restriction enzyme digestion, by PCR amplification, etc. For the most part, DNA fragments will be of at least 15 nt, usually at least 18 nt, more usually at least about 50 nt. Such small DNA fragments are useful as primers for PCR, hybridization screening, etc. Larger DNA fragments, i.e. greater than 100 nt are useful for production of the encoded polypeptide. For use in amplification reactions, such as PCR, a pair of primers will be used. The exact composition of the primer sequences is not critical to the invention, but for most applications the primers will hybridize to the subject sequence under stringent conditions, as known in the art. It is preferable to chose a pair of primers that will generate an amplification product of at least about 50 nt, preferably at least about 100 nt. Algorithms for the selection of primer sequences are generally known, and are available in commercial software packages.

Amplification primers hybridize to complementary strands of DNA, and will prime towards each other.

The *ptc* genes are isolated and obtained in substantial purity, generally as other than an intact mammalian chromosome. Usually, the DNA will be obtained substantially free of other nucleic acid sequences that do not include a *ptc* sequence or fragment thereof, generally being at least about 50%, usually at least about 90% pure and are typically "recombinant", i.e. flanked by one or more nucleotides with which it is not normally associated on a naturally occurring chromosome.

The DNA sequences are used in a variety of ways. They may be used as probes for identifying other patched genes. Mammalian homologs have substantial sequence similarity to the subject sequences, i.e. at least 75%, usually at least 90%, more usually at least 95% sequence identity with the nucleotide sequence of the subject DNA sequence. Sequence similarity is calculated based on a reference sequence, which may be a subset of a larger sequence, such as a conserved motif, coding region, flanking region, etc. A reference sequence will usually be at least about 18 nt long, more usually at least about 30 nt long, and may extend to the complete sequence that is being compared. Algorithms for sequence analysis are known in the art, such as BLAST, described in Altschul *et al.* (1990) J Mol Biol 215:403-10.

Nucleic acids having sequence similarity are detected by hybridization under low stringency conditions, for example, at 50 C and 10xSSC (0-9 M saline/0.09 M sodium citrate) and remain bound when subjected to washing at 55°C in 1xSSC. By using probes, particularly labeled probes of DNA sequences, one can isolate homologous or related genes. The source of homologous genes may be any mammalian species, e.g. primate species, particularly human- murines, such as rats and mice, canines, felines, bovines, ovines, equines, etc.

The DNA may also be used to identify expression of the gene in a biological specimen. The manner in which one probes cells for the presence of particular nucleotide sequences, as genomic DNA or RNA, is well-established in the literature and does not require elaboration here. Conveniently, a biological specimen is used as a source of mRNA. The mRNA may be amplified by RT-PCR, using reverse transcriptase to form a complementary DNA strand, followed by polymerase chain reaction amplification using primers specific for the subject DNA sequences. Alternatively, the mRNA sample is separated by gel electrophoresis, transferred to a suitable support, e.g., nitrocellulose and then probed with a fragment of the subject DNA as a probe. Other techniques may also find use. Detection of mRNA having the subject sequence is indicative of *patched* gene expression in the sample.

The subject nucleic acid sequences may be modified for a number of purposes, particularly where they will be used intracellularly, for example, by being joined to a nucleic acid cleaving agent, e.g. a chelated metal ion, such as iron or chromium for cleavage of the

gene; as an antisense sequence, or the like. Modifications may include replacing oxygen of the phosphate esters with sulfur or nitrogen, replacing the phosphate with phosphoramide, etc.

A number of methods are available for analyzing genomic DNA sequences. Where large amounts of DNA are available, the genomic DNA is used directly. Alternatively, the region of interest is cloned into a suitable vector and grown in sufficient quantity for analysis, or amplified by conventional techniques, such as the polymerase chain reaction (PCR). The use of the polymerase chain reaction is described in Saiki, *et al.* (1985) Science 239:487, and a review of current techniques may be found in Sambrook, *et al.* Molecular Cloning: A Laboratory Manual, CSH Press 1989, pp.14.2-14.33.

A detectable label may be included in the amplification reaction. Suitable labels include fluorochromes, e.g. fluorescein isothiocyanate (FITC), rhodamine, Texas Red, phycoerythrin, allophycocyanin, 6-carboxyfluorescein (6-FAM), 2',7'-dimethoxy-4',5'-dichloro-6-carboxyfluorescein (JOE), 6-carboxy-Xrhodamine (ROX), 6-carboxy-2',4',7',4,7-hexachlorofluorescein (HEX), 5-carboxyfluorescein (5-FAM) or N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA), radioactive labels, e.g. **32P, 35S,** 3H; etc. The label may be a two stage system, where the amplified DNA is conjugated to biotin, haptens, etc. having a high affinity binding partner, e.g. avidin, specific antibodies, etc., where the binding partner is conjugated to a detectable label. The label may be conjugated to one or both of the primers. Alternatively, the pool of nucleotides used in the amplification is labeled, so as to incorporate the label Into the amplification product.

The amplified or cloned fragment may be sequenced by dideoxy or other methods, and the sequence of bases compared to the normal *ptc* sequence. Hybridization with the variant sequence may also be used to determine its presence, by Southern blots, dot blots, etc. Single strand conformational polymorphism (SSCP) analysis, denaturing gradient gel electrophoresis (DGGE), and heteroduplex analysis in gel matrices are used to detect conformational changes created by DNA sequence variation as alterations in electrophoretic mobility. The hybridization pattern of a control and variant sequence to an array of oligonucleotide probes immobilized on a solid support, as described in WO 95/11995, may also be used as a means of detecting the presence of variant sequences. Alternatively, where a predisposing mutation creates or destroys a recognition site for a restriction endonuclease, the fragment is digested with that endonuclease, and the products size fractionated to determine whether the fragment was digested. Fractionation is performed by gel electrophoresis, particularly acrylamide or agarose gels.

In a merely illustrative embodiment, the method includes the steps of (i) collecting a sample of cells from a patient, (ii) isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, (iii) contacting the nucleic acid sample with one or more primers

which specifically hybridize to a *ptc* gene under conditions such that hybridization and amplification of the *ptc* gene (if present) occurs, and (iv) detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample.

In yet another exemplary embodiment, aberrant methylation patterns of a *ptc* gene can be detected by digesting genomic DNA from a patient sample with one or more restriction endonucleases that are sensitive to methylation and for which recognition sites exist in the *ptc* gene (including in the flanking and intronic sequences). See, for example, Buiting et al., (1994) Human Mol Genet 3:893-895. Digested DNA is separated by gel electrophoresis, and hybridized with probes derived from, for example, genomic or cDNA sequences. The methylation status of the *ptc* gene can be determined by comparison of the restriction pattern generated from the sample DNA with that for a standard of known methylation.

In still another embodiment, a diagnostic assay is provided which detects the ability of a *ptc* gene product, e.g., recombinantly expressed from a gene isolated from a biopsied cell, to bind to other proteins, e.g., upstream (*hedgehog*) or downstream of *ptc*. For instance, it will be desirable to detect *ptc* mutants which bind with lower binding affinity for *hedgehog* proteins. Such mutants may arise, for example, from fine mutations, e.g., point mutants, which may be impractical to detect by the diagnostic DNA sequencing techniques or by the immunoassays described above. The present invention accordingly further contemplates diagnostic screening assays which generally comprise cloning one or more *ptc* genes from the sample cells, and expressing the cloned genes under conditions which permit detection of an interaction between that recombinant gene product and a *ptc*-binding protein, e.g., a *hedgehog* protein. As will be apparent from the description of the various drug screening assays set forth below, a wide variety of techniques can be used to determine the ability of a *ptc* protein to bind to other cellular components.

The subject nucleic acids can be used to generate transgenic animals or site specific gene modifications in cell lines. Transgenic animals may be made through homologous recombination, where the normal patched locus is altered. Alternatively, a nucleic acid construct is randomly integrated into the genome, Vectors for stable integration include plasmids, retroviruses and other animal viruses, YACS, and the like.

The modified cells or animals are useful in the study of *patched* function and regulation. For example, a series of small deletions and/or substitutions may be made in the *patched* gene to determine the role of different exons in oncogenesis, signal transduction, etc. Of particular interest are transgenic animal models for carcinomas of the skin, where expression of *ptc* is specifically reduced or absent in skin cells. An alternative approach to transgenic models for this disease are those where one of the mammalian *hedgehog* genes, *e.g. Shh, Ihh, Dhh,* are upregulated in skin cells, or in other cell types. For models of skin

abnormalities, one may use a skin-specific promoter to drive expression of the transgene, or other inducible promoter that can be regulated in the animal model. Such promoters include keratin gene promoters. Specific constructs of interest include anti-sense *ptc*, which will block *ptc* expression, expression of dominant negative *ptc* mutations, and over-expression of HH genes. A detectable marker, such as *lacZ* may be introduced into the *patched* locus, where upregulation of patched expression will result in an easily detected change in phenotype.

One may also provide for expression of the *patched* gene or variants thereof in cells or tissues where it is not normally expressed or at abnormal times of development. Thus, mouse models of spina bifida or abnormal motor neuron differentiation in the developing spinal cord are made available. In addition, by providing expression of *ptc* protein in cells in which it is otherwise not normally produced, one can induce changes in cell behavior, e.g. through *ptc* mediated transcription modulation.

DNA constructs for homologous recombination will comprise at least a portion of the *patched or hedgehog* gene with the desired genetic modification, and will include regions of homology to the target locus. DNA constructs for random integration need not include regions of homology to mediate recombination. Conveniently, markers for positive and negative selection are included. Methods for generating cells having targeted gene modifications through homologous recombination are known in the art. For various techniques for transfecting mammalian cells, see Keown *et al.* (1990) Methods in Enzymology 185:527-537.

For embryonic stem (ES) cells, an ES cell line may be employed, or ES cells may be obtained freshly from a host, e.g. mouse, rat, guinea pig, etc. Such cells are grown on an appropriate fibroblast-feeder layer or grown in the presence of leukemia inhibiting factor (LIF). When ES cells have been transformed, they may be used to produce transgenic animals. After transformation, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be detected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination or integration of the construct. Those colonies that are positive may then be used for embryo manipulation and blastocyst injection. Blastocysts are obtained from 4 to 6 week old superovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant females. Females are then allowed to go to term and the resulting litters screened for mutant cells having the construct. By providing for a different phenotype of the blastocyst and the ES cells, chimeric progeny can be readily detected.

The chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogeneic or congenic grafts or transplants, or in *in vitro* culture. The transgenic animals may be any non-human mammal, such as laboratory animals, domestic animals, etc. The transgenic animals may be used in functional studies, drug screening, etc., e.g. to determine the effect of a candidate drug on basal cell carcinomas.

The subject gene may be employed for producing all or portions of the patched protein. For expression, an expression cassette may be employed, providing for a transcriptional and translational initiation region, which may be inducible or constitutive, the coding region under the transcriptional control of the transcriptional initiation region, and a transcriptional and translational termination region. Various transcriptional initiation regions may be employed which are functional in the expression host.

Specific *ptc* peptides of interest include the extracellular domains, particularly in the human mature protein, aa 120 to 437, and aa 770 to 1027. These peptides may be used as immunogens to raise antibodies that recognize the protein in an intact cell membrane. The cytoplasmic domains, as shown in Figure 2, (the amino terminus and carboxy terminus) are of interest in binding assays to detect ligands involved in signaling mediated by *ptc*.

The peptide may be expressed in prokaryotes or eukaryotes in accordance with conventional ways, depending upon the purpose for expression. For large scale production of the protein, a unicellular organism or cells of a higher organism, e.g. eukaryotes such as vertebrates, particularly mammals, may be used as the expression host, such as *E. coli, B,* subtilis ~~*subthis*~~, *S. cerevisiae,* and the like. In many situations, it may be desirable to express the *patched* gene in a mammalian host, whereby the *patched* gene will be glycosylated, and transported to the cellular membrane for various studies.

RS 5/8/02

With the availability of the protein in large amounts by employing an expression host, the protein may be isolated and purified in accordance with conventional ways. A lysate may be prepared of the expression host and the lysate purified using HPLC, exclusion chromatography, gel electrophoresis, affinity chromatography, or other purification technique. The purified protein will generally be at least about 80% pure, preferably at least about 90% pure, and may be up to and including 100% pure. By pure is intended free of other proteins, as well as cellular debris.

The polypeptide is used for the production of antibodies, where short fragments provide for antibodies specific for the particular polypeptide, whereas larger fragments or the entire gene allow for the production of antibodies over the surface of the polypeptide or protein. Antibodies may be raised to the normal or mutated forms of *ptc*- The extracellular domains of the protein are of interest as epitopes, particular antibodies that recognize

common changes found in abnormal, oncogenic *ptc*, which compromise the protein activity. Antibodies may be raised to isolated peptides corresponding to these domains, or to the native protein, e.g. by immunization with cells expressing *ptc*, immunization with liposomes having *ptc* inserted in the membrane, etc. Antibodies that recognize the extracellular domains of *ptc* are useful in diagnosis, typing and staging of human carcinomas.

Antibodies are prepared in accordance with conventional ways, where the expressed polypeptide or protein may be used as an immunogen, by itself or conjugated to known immunogenic carriers, e.g. KLH, pre-S HBsAg, other viral or eukaryotic proteins, or the like. Various adjuvants may be employed, with a series of injections, as appropriate, For monoclonal antibodies, after one or more booster injections, the spleen may be isolated, the splenocytes immortalized, and then screened for high affinity antibody binding. The immortalized cells, e.g. hybridomas, producing the desired antibodies may then be expanded. For further description, see Monoclonal Antibodies- A Laboratory Manual, Harlow and Lane eds., Cold Spring Harbor Laboratories, Cold Spring Harbor, New York, 1988. If desired, the mRNA encoding the heavy and light chains may be isolated and mutagenized by cloning in *E. coli,* and the heavy and light chains may be mixed to further enhance the affinity of the antibody.

The antibodies find particular use in diagnostic assays for developmental abnormalities, basal cell carcinomas and other tumors associated with mutations in *ptc*. Staging, detection and typing of tumors may utilize a quantitative immunoassay for the presence or absence of normal *ptc*. Alternatively, the presence of mutated forms of *ptc* may be determined. A reduction in normal *ptc* and/or presence of abnormal *ptc* is indicative that the tumor is *ptc*-associated.

A sample is taken from a patient suspected of having a *ptc*-associated tumor, developmental abnormality or BCNS. Samples, as used herein, include biological fluids such as blood, cerebrospinal fluid, tears, saliva, lymph, dialysis fluid and the like- organ or tissue culture derived fluids, and fluids extracted from physiological tissues. Also included in the term are derivatives and fractions of such fluids. Biopsy samples are of particular interest, e.g. skin lesions, organ tissue fragments, etc. Where metastasis is suspected, blood samples may be preferred. The number of cells in a sample will generally be at least about 10^3, usually at least 10^4 more usually at least about 10^5. The cells may be dissociated, in the case of solid tissues, or tissue sections may be analyzed. Alternatively a lysate of the cells may be prepared.

Diagnosis may be performed by a number of methods. The different methods all determine the absence or presence of normal or abnormal *ptc* in patient cells suspected of having a mutation in *ptc*. For example, detection may utilize staining of intact cells or histological sections, performed in accordance with conventional methods. The antibodies of

interest are added to the cell sample, and incubated for a period of time sufficient to allow binding to the epitope, usually at least about 10 minutes. The antibody may be labeled with radioisotopes, enzymes, fluorescers, chemiluminescers, or other labels for direct detection. Alternatively, a second stage antibody or reagent is used to amplify the signal. Such reagents are well-known in the art. For example, the primary antibody may be conjugated to biotin, with horseradish peroxidase-conjugated avidin added as a second stage reagent. Final detection uses a substrate that undergoes a color change in the presence of the peroxidase. The absence or presence of antibody binding may be determined by various methods, including flow cytometry of dissociated cells, microscopy, radiography, scintillation counting, etc.

An alternative method for diagnosis depends on the *in vitro* detection of binding between antibodies and *ptc* in a lysate. Measuring the concentration of *ptc* binding in a sample or fraction thereof may be accomplished by a variety of specific assays. A conventional sandwich type assay may be used. For example, a sandwich assay may first attach *ptc*-specific antibodies to an insoluble surface or support. The particular manner of binding is not crucial so long as it is compatible with the reagents and overall methods of the invention They may be bound to the plates covalently or non-covalently, preferably non-covalently.

The insoluble supports may be any compositions to which polypeptides can be bound, which is readily separated from soluble material, and which is otherwise compatible with the overall method. The surface of such supports may be solid or porous and of any convenient shape. Examples of suitable insoluble supports to which the receptor is bound include beads, e.g. magnetic beads, membranes and microtiter plates. These are typically made of glass, plastic (e.g. polystyrene), polysaccharides, nylon or nitrocellulose. Microtiter plates are especially convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples.

Patient sample lysates are then added to separately assayable supports (for example, separate wells of a microtiter plate) containing antibodies. Preferably, a series of standards, containing known concentrations of normal and/or abnormal *ptc* is assayed in parallel with the samples or aliquots thereof to serve as controls. Preferably, each sample and standard will be added to multiple wells so that mean values can be obtained for each. The incubation time should be sufficient for binding, generally, from about 0.1 to 3 hr is sufficient. After incubation, the insoluble support is generally washed of non-bound components. Generally, a dilute non-ionic detergent medium at an appropriate pH, generally 7-8, is used as a wash medium. From one to six washes may be employed, with sufficient volume to thoroughly wash nonspecifically bound proteins present in the sample.

After washing, a solution containing a second antibody is applied. The antibody will bind *ptc* with sufficient specificity such that it can be distinguished from other components present. The second antibodies may be labeled to facilitate direct, or indirect quantification of binding. Examples of labels that permit direct measurement of second receptor binding include radiolabels, such as ^{3}H or ^{125}I, fluorescers, dyes, beads, chemilumninescers, colloidal particles, and the like. Examples of labels which permit indirect measurement of binding include enzymes where the substrate may provide for a colored or fluorescent product. In a preferred embodiment, the antibodies are labeled with a covalently bound enzyme capable of providing a detectable product signal after addition of suitable substrate. Examples of suitable enzymes for use in conjugates include horseradish peroxidase, alkaline phosphatase, malate dehydrogenase and the like. Where not commercially available, such antibody-enzyme conjugates are readily produced by techniques known to those skilled in the art. The incubation time should be sufficient for the labeled ligand to bind available molecules. Generally, from about 0. 1 to 3 hr is sufficient, usually 1 hr sufficing.

After the second binding step, the insoluble support is again washed free of non-specifically bound material. The signal produced by the bound conjugate is detected by conventional means. Where an enzyme conjugate is used, an appropriate enzyme substrate is provided so a detectable product is formed.

Other immunoassays are known in the art and may find use as diagnostics. Ouchterlony plates provide a simple determination of antibody binding. Western blots may be performed on protein gels or protein spots on filters, using a detection system specific for *ptc* as desired, conveniently using a labeling method as described for the sandwich assay.

Other diagnostic assays of interest are based on the functional properties of *ptc* protein itself. Such assays are particularly useful where a large number of different sequence changes lead to a common phenotype, i.e., loss of protein function leading to oncogenesis or developmental abnormality. For example, a functional assay may be based on the transcriptional changes mediated by *hedgehog* and *patched* gene products. Addition of soluble Hh to embryonic stem cells causes induction of transcription in target genes. The presence of functional *ptc* can be determined by its ability to antagonize Hh activity. Other functional assays may detect the transport of specific molecules mediated by *ptc*, in an intact cell or membrane fragment. Conveniently, a labeled substrate is used, where the transport in or out of the cell can be quantitated by radiography, microscopy, flow cytometry, spectrophotometry, etc. Other assays may detect conformational changes, or changes in the subcellular localization of *patched* protein.

By providing for the production of large amounts of patched protein, one can identify ligands or substrates that bind to, modulate or mimic the action of *patched*. A common feature in basal cell carcinoma is the loss of adhesion between epidermal and dermal layers,

indicating a role for *ptc* in maintaining appropriate cell adhesion. Areas of investigation include the development of cancer treatments, wound healing, adverse effects of aging, metastasis, etc.

Drug screening identifies agents that provide a replacement for *ptc* function in abnormal cells. The role of *ptc* as a tumor suppressor indicates that agents which mimic its function, in terms of transmembrane transport of molecules, transcriptional down-regulation, etc., will inhibit the process of oncogenesis. These agents may also promote appropriate cell adhesion in wound healing and aging, to reverse the loss of adhesion observed in metastasis, etc. Conversely, agents that reverse *ptc* function may stimulate controlled growth and healing. Of particular interest are screening assays for agents that have a low toxicity for human cells. A wide variety of assays may be used for this purpose, including labeled *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays for protein binding, and the like. The purified protein may also be used for determination of three-dimensional crystal structure, which can be used for modeling intermolecular interactions, transporter function, etc.

The term "agent" as used herein describes any molecule, e.g. protein or pharmaceutical, with the capability of altering or mimicking the physiological function of *patched*. Generally a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e. at zero concentration or below the level of detection.

Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or a combinations thereof.

Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides and oligopeptides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are

readily modified through conventional chemical, physical and biochemical means, and may be used to produce combinatorial libraries. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs.

Where the screening assay is a binding assay, one or more of the molecules may be joined to a label, where the label can directly or indirectly provide a detectable signal. Various labels include radioisotopes, fluorescers, chemiluminescers, enzymes, specific binding molecules, particles, e.g. magnetic particles, and the like. Specific binding molecules include pairs, such as biotin and streptavidin, digoxin and antidigoxin etc. For the specific binding members, the complementary member would normally be labeled with a molecule that provides for detection, in accordance with known procedures.

A variety of other reagents may be included in the screening assay. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc. that are used to facilitate optimal protein-protein binding and/or reduce nonspecific or background interactions. Reagents that improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc. may be used. The mixture of components are added in any order that provides for the requisite binding. Incubations are performed at any suitable temperature, typically between 4° and 40° C. Incubation periods are selected for optimum activity, but may also be optimized to facilitate rapid high-throughput screening. Typically between 0.1 and 1 hours will be sufficient.

Other assays of interest detect agents that mimic patched function, such as repression of target gene transcription, transport of patched substrate compounds, etc. For example, an expression construct comprising a *patched* gene may be introduced into a cell line under conditions that allow expression. The level of patched activity is determined by a functional assay, as previously described. In one screening assay, candidate agents are added in combination with a Hh protein, and the ability to overcome Hh antagonism of *ptc* is detected. In another assay, the ability of candidate agents to enhance *ptc* function is determined. Alternatively, candidate agents are added to a cell that lacks functional *ptc*, and screened for the ability to reproduce *ptc* in a functional assay.

In one embodiment, the drug screening assay is a cell-based assay which detects the ability of a compound to alter *patched*-dependent gene transcription. By selecting transcriptional regulatory sequences from genes whose expression is regulated by *patched* signal transduction, e.g. from *patched, GLI, hedgehog* or PTHrP genes, e.g., regulatory sequences that are responsible for the up- or down regulation of these genes in response to *patched* signalling, and operatively linking such promoters to a reporter gene, one can derive a transcription based assay which is sensitive to the ability of a specific test compound to modify *patched* signalling pathways. Expression of the reporter gene, thus, provides a

valuable screening tool for the development of compounds that act as agonists or antagonists of *patched*.

Reporter gene based assays of this invention measure the end stage of the above described cascade of events, e.g., transcriptional modulation. Accordingly, in practicing one embodiment of the assay, a reporter gene construct is inserted into the reagent cell in order to generate a detection signal dependent on *ptc* signaling. To identify potential regulatory elements responsive to *ptc* signaling present in the transcriptional regulatory sequence of a target gene, nested deletions of genomic clones of the target gene can be constructed using standard techniques. See, for example, Current Protocols in Molecular Biology, Ausubel, F.M. et al. (eds.) Greene Publishing Associates, (1989); U.S. Patent 5,266,488; Sato et al. (1995) *J Biol Chem* 270:10314-10322; and Kube et al. (1995) *Cytokine* 7:1-7. A nested set of DNA fragments from the gene's 5'-flanking region are placed upstream of a reporter gene, such as the luciferase gene, and assayed for their ability to direct reporter gene expression in *patched* expressing cells. Host cells transiently transfected with reporter gene constructs can be scored for the induction of expression of the reporter gene in the presence and absence of *hedgehog* to determine regulatory sequences which are responsive to *patched*-dependent signalling.

RS 5/10/02

In practicing one embodiment of the assay, a reporter gene construct is inserted into the reagent cell in order to generate a detection signal dependent on second messengers generated by induction with *hedgehog* protein. Typically, the reporter gene construct will include a reporter gene in operative linkage with one or more transcriptional regulatory elements responsive to the *hedgehog* activity, with the level of expression of the reporter gene providing the *hedgehog*-dependent detection signal. The amount of transcription from the reporter gene may be measured using any method known to those of skill in the art to be suitable. For example, mRNA expression from the reporter gene may be detected using RNAse protection or RNA-based PCR, or the protein product of the reporter gene may be identified by a characteristic stain or an intrinsic activity. The amount of expression from the reporter gene is then compared to the amount of expression in either the same cell in the absence of the test compound (or *hedgehog*) or it may be compared with the amount of transcription in a substantially identical cell that lacks the target receptor protein. Any statistically or otherwise significant difference in the amount of transcription indicates that the test compound has in some manner altered the signal transduction of the *patched* protein, e.g., the test compound is a potential *ptc* therapeutic.

As described in further detail below, in preferred embodiments the gene product of the reporter is detected by an intrinsic activity associated with that product. For instance, the reporter gene may encode a gene product that, by enzymatic activity, gives rise to a detection signal based on color, fluorescence, or luminescence. In other preferred embodiments, the

reporter or marker gene provides a selective growth advantage, e.g., the reporter gene may enhance cell viability, relieve a cell nutritional requirement, and/or provide resistance to a drug.

Preferred reporter genes are those that are readily detectable. The reporter gene may also be included in the construct in the form of a fusion gene with a gene that includes desired transcriptional regulatory sequences or exhibits other desirable properties. Examples of reporter genes include, but are not limited to CAT (chloramphenicol acetyl transferase) (Alton and Vapnek (1979), Nature 282: 864-869) luciferase, and other enzyme detection systems, such as beta-galactosidase; firefly luciferase (deWet et al. (1987), Mol. Cell. Biol. 7:725-737); bacterial luciferase (Engebrecht and Silverman (1984), PNAS 1: 4154-4158; Baldwin et al. (1984), Biochemistry 23: 3663-3667); alkaline phosphatase (Toh et al. (1989) Eur. J. Biochem. 182: 231-238, Hall et al. (1983) J. Mol. Appl. Gen. 2: 101), human placental secreted alkaline phosphatase (Cullen and Malim (1992) Methods in Enzymol. 216:362-368).

Transcriptional control elements which may be included in a reporter gene construct include, but are not limited to, promoters, enhancers, and repressor and activator binding sites. Suitable transcriptional regulatory elements may be derived from the transcriptional regulatory regions of genes whose expression is induced after modulation of a *patched* signal transduction pathway. The characteristics of preferred genes from which the transcriptional control elements are derived include, but are not limited to, low or undetectable expression in quiescent cells, rapid induction at the transcriptional level within minutes of extracellular simulation, induction that is transient and independent of new protein synthesis, subsequent shut-off of transcription requires new protein synthesis, and mRNAs transcribed from these genes have a short half-life. It is not necessary for all of these properties to be present.

The compounds having the desired pharmacological activity may be administered in a physiologically acceptable carrier to a host for treatment of cancer or developmental abnormalities attributable to a defect in *patched* function. The compounds may also be used to enhance *patched* function in wound healing, aging, etc. The inhibitory agents may be administered in a variety of ways, orally, topically, parenterally e.g. subcutaneously, intraperitoneally, by viral infection, intravascularly, etc. Topical treatments are of particular interest. Depending upon the manner of introduction, the compounds may be formulated in a variety of ways. The concentration of therapeutically active compound in the formulation may vary from about 0.1-100 wt.%.

The pharmaceutical compositions can be prepared in various forms, such as granules, tablets, pills, suppositories, capsules, suspensions, salves, lotions and the like. Pharmaceutical grade organic or inorganic carriers and/or diluents suitable for oral and topical use can be used to make up compositions containing the therapeutically-active compounds. Diluents known to the art include aqueous media, vegetable and animal oils and fats. Stabilizing

agents, wetting and emulsifying agents, salts for varying the osmotic pressure or buffers for securing an adequate pH value, and skin penetration enhancers can be used as auxiliary agents.

The gene or fragments thereof may be used as probes for identifying the 5' non-coding region comprising the transcriptional initiation region, particularly the enhancer regulating the transcription of *patched.* By probing a genomic library, particularly with a probe comprising the 5' coding region, one can obtain fragments comprising the 5' non-coding region. If necessary, one may walk the fragment to obtain further 5' sequence to ensure that one has at least a functional portion of the enhancer. It is found that the enhancer is proximal to the 5' coding region, a portion being in the transcribed sequence and downstream from the promoter sequences. The transcriptional initiation region may be used for many purposes, studying embryonic development, providing for regulated expression of *patched* protein or other protein of interest during embryonic development or thereafter, and in gene therapy.

The gene may also be used for gene therapy. Vectors useful for introduction of the gene include plasmids and viral vectors. Of particular interest are retroviral-based vectors, e.g. moloney murine leukemia virus and modified human immunodeficiency virus-adenovirus vectors, etc. Gene therapy may be used to treat skin lesions, an affected fetus, etc., by transfection of the normal gene into embryonic stem cells or into other fetal cells. A wide variety of viral vectors can be employed for transfection and stable integration of the gene into the genome of the cells. Alternatively, micro-injection may be employed, fusion, or the like for introduction of genes into a suitable host cell. See, for example, Dhawan *et al.* (1991) Science 254:1509-1512 and Smith *et al.* (1 990) Molecular and Cellular Biology 3268-3271.

The following examples are offered by illustration not by way of limitation.

EXPERIMENTAL

Methods and Materials

PCR on Mosquito (Anopheles gambiae) Genomic DNA. PCR primers were based on amino acid stretches of fly *ptc* that were not likely to diverge over evolutionary time and were of low degeneracy. Two such primers (P2Rl (SEO ID NO-14)-GGACGAATTCAARGTNCAYCARYTNTGG, P4Rl: (SEQ ID NO:15) GGACGAATTCCYTCCCARAARCANTC, (the underlined sequences are Eco RI linkers) amplified an appropriately sized band from mosquito genomic DNA using the PCR. The program conditions were as follows:

94 C 4 min.; 72 C Add Taq;

[49 C 30 sec.; 72 C 90 sec.; 94 C 15 sec] 3 times

[94 C 15 sec.; 50 C 30 sec.; 72 C 90 sec] 35 times

72 C 10 min; 4 C hold

This band was subcloned into the EcoRV site of pBluescript II and sequenced using the USB Sequence kit.

Screen of a Butterfly cDNA Library with Mosquito PCR Product. Using the mosquito PCR product (SEQ ID NO:7) as a probe, a 3 day embryonic *Precis coenia* gt10 cDNA library (generously provided by Sean Carroll) was screened. Filters were hybridized at 65° C overnight in a solution containing 5xSSC, 10% dextran sulfate, 5x Denhardt's, 200 μg/ml sonicated salmon sperm DNA, and 0.5% SDS. Filters were washed in 0.1X SSC, 0.1% SDS at room temperature several times to remove nonspecific hybridization. Of the 100,000 plaques initially screened, 2 overlapping clones, Ll and L2, were isolated, which corresponded to the N terminus of butterfly *ptc*. Using L2 as a probe, the library filters were rescreened and 3 additional clones (L5, L7, L8) were isolated which encompassed the remainder of the *ptc* coding sequence. The full length sequence of butterfly *ptc* (SEQ ID NO:3) was determined by ABI automated sequencing.

Screen of a Tribolium (beetle) Genomic Library with Mosquito PCR Product and 900 bp Fragment from the Butterfly Clone. A gem11 genomic library from *Tribolium casteneum* (gift of Rob Dennell) was probed with a mixture of the mosquito PCR (SEQ ID NO:7) product and BstXI/EcoRI fragment of L2. Filters were hybridized at 55 C overnight and washed as above. Of the 75,000 plaques screened, 14 clones were identified and the Sacl fragment of T8 (SEQ ID NO:1), which crosshybridized with the mosquito and butterfly probes, was subcloned into pBluescript.

PCR on Mouse cDNA Using Degenerate Primers Derived from Regions Conserved in the Four Insect Homologues. Two degenerate PCR primers (P4REV- (SEQ ID NO:16) <u>GGACGAATTC</u>YTNGANTGYTTYTGGGA- P22- (SEQ ID NO:17) <u>CATACCAGCCAAG CTTG</u>TCIGGCCARTGCAT) were designed based on a comparison of *ptc* amino acid sequences from fly *(Drosophila melanogaster)* (SEQ ID NO:6), mosquito *(Anopheles gambiae)* (SEQ ID NO:8), butterfly (*Precis coenia*) (SEQ ID NO:4), and beetle *(Tribolium casteneum)* (SEQ ID NO:2). I represents inosine, which can form base pairs with all four nucleotides. P22 was used to reverse transcribe RNA from 12.5 dpc mouse limb bud (gift from David Kingsley) for 90 min at 37 C. PCR using P4REV (SEQ ID NO:17) and P22 (SEQ ID NO:18) was then performed on 1 l of the resultant cDNA under the following conditions:

94 C 4 min.; 72 C Add Taq;

[94 C 15 sec.- 50 C 30 sec.- 72 C 90 sec.] 35 times

72 C 10 min.-, 4 C hold

PCR products of the expected size were subcloned into the TA vector (Invitrogen) and sequenced with the Sequenase Version 2.0 DNA Sequencing Kit (U. S. B.).

Using the cloned mouse PCR fragment as a probe, 300,000 plaques of a mouse 8.5 dpc gtl0 cDNA library (a gift from Brigid Hogan) were screened at 65 C as above and washed in 2x SSC, 0.1% SDS at room temperature. 7 clones were isolated, and three (M2, M4, and M8) were subcloned into pBluescript II. 200,000 plaques of this library were rescreened using first, a 1.1 kb EcoRI fragment from M2 to identify 6 clones (M9-Ml6) and secondly a mixed probe containing the most N terminal (Xhol fragment from M2) and most C terminal sequences (BamHI/BglII fragment from M9) to isolate 5 clones (Ml7-M21). M9, M10, M14, and M17-21 were subcloned into the EcoRI site of pBluescript II (Strategene).

RNA Blots and in situ Hybridizations in Whole and Sectioned Mouse Embryos:

Northerns. A mouse embryonic Northern blot and an adult multiple tissue Northern blot (obtained from Clontech) were probed with a 900 bp EcoRl fragment from an N terminal coding region of mouse *ptc*. Hybridization was performed at 65° C in 5x SSPE, l0x Denhardt's, 100 μg/ml sonicated salmon sperm DNA, and 2% SDS. After several short room temperature washes in 2x SSC, 0.05% SDS, the blots were washed at high stringency in 0. 1 X SSC, 0.1% SDS at 50° C.

In situ hybridization of sections: 7.75, 8.5, 11.5, and 13.5 dpc mouse embryos were dissected in PBS and frozen in Tissue-Tek medium at -80° C. 12-16 μm frozen sections were cut, collected onto VectaBond (Vector Laboratories) coated slides, and dried for 30-60 minutes at room temperature. After a 10 minute fixation in 4% paraformaldehyde in PBS, the slides were washed 3 times for 3 minutes in PBS, acetylated for 10 minutes in 0.25% acetic anhydride in triethanolamine, and washed three more times for 5 minutes in PBS. Prehybridization (50% formamide, 5X SSC, 250 μg/ml yeast tRNA, 500 μg/ml sonicated salmon sperm DNA, and 5x Denhardt's) was carried out for 6 hours at room temperature in 50% formamide/5x SSC humidified chambers. The probe, which consisted of 1 kb from the N-terminus of *ptc*, was added at a concentration of 200-1000 ng/ml into the same solution used for prehybridization, and then denatured for five minutes at 80° C. Approximately 75 μl of probe were added to each slide and covered with Parafilm. The slides were incubated overnight at 65° C in the same humidified chamber used previously. The following day, the probe was washed successively in 5X SSC (5 minutes, 65° C), 0.2X SSC (1 hour, 65° C), and 0.2X SSC (10 minutes, room temperature). After five minutes in buffer Bl (0.lM maleic acid, 0.15 M NaCl, pH 7.5), the slides were blocked for 1 hour at room temperature in 1% blocking reagent (Boerhinger-Mannheim) in buffer Bl, and then incubated for 4 hours in buffer Bl containing the DIG-AP conjugated antibody (Boerhinger-Mannheim) at a 1:5000 dilution. Excess antibody was removed during two 15 minute washes in buffer Bl, followed by five

minutes in buffer B3 (100 mM Tris, l00mM NaCl, 5mM MgCl2, pH 9.5). The antibody was detected by adding an alkaline phosphatase substrate (350 μl 75 mg/ml X-phosphate in DMF, 450 μl 50 mg/ml NBT in 70% DMF in 100 mls of buffer B3) and allowing the reaction to proceed overnight in the dark. After a brief rinse in 10 mM Tris, 1mM EDTA, pH 8.0, the slides were mounted with Aquamount (Lerner Laboratories).

Drosophila 5-transcriptional initiation region -gal constructs. A series of constructs were designed that link different regions of the *ptc* promoter from *Drosophila* to a LacZ reporter gene in order to study the cis regulation of the *ptc* expression pattern. See Fig. 1. A 10.8kb BamHI/BspMl fragment comprising the 5'-non-coding region of the mRNA at its 3'-terminus was obtained and truncated by restriction enzyme digestion as shown in Fig. 1. These expression cassettes were introduced into *Drosophila* lines using a P-element vector (Thummel *et al.* (1988) Gene 74:445-456), which were injected into embryos, providing flies which could be grown to produce embryos. (See Spradling and Rubin (1982) Science 218:341-347 for a description of the procedure.) The vector used a pUC8 background into which was introduced the white gene to provide for yellow eyes, portions of the P-element for integration, and the constructs were inserted into a polylinker upstream from the LacZ gene. The resulting embryos, larvae, and adults were stained using antibodies to LacZ protein conjugated to HRP and the samples developed with OPD dye to identify the expression of the LacZ gene. The staining pattern in embryos is described in Fig. 1, indicating whether there was staining during the early and late development of the embryo.

Isolation of a Mouse ptc Gene. Homologues of fly *ptc* (SEQ ID NO:6) were isolated from three insects: mosquito, butterfly and beetle, using either PCR or low stringency library screens. PCR primers to six amino acid stretches of *ptc* of low mutatability and degeneracy were designed. One primer pair, P2 and P4, amplified an homologous fragment of *ptc* from mosquito genomic DNA that corresponded to the first hydrophilic loop of the protein. The 345bp PCR product (SEQ ID NO:7) was subcloned and sequenced and when aligned to fly *ptc*, showed 67% amino acid identity.

The cloned mosquito fragment was used to screen a butterfly gt 10 cDNA library. Of 100,000 plaques screened, five overlapping clones were isolated and used to obtain the full length coding sequence. The butterfly *ptc* homologue (SEQ ID NO:4) is 1,311 amino acids long and overall has 50% amino acid identity (72% similarity) to fly *ptc*. With the exception of a divergent C-terminus, this homology is evenly spread across the coding sequence. The mosquito PCR clone (SEQ ID NO:7) and a corresponding fragment of butterfly cDNA were used to screen a beetle gemll genomic library. Of the plaques screened, 14 clones were identified. A fragment of one clone (T8), which hybridized with the original probes, was subcloned and sequenced. This 3kb piece contains an 89 amino acid exon (SEQ ID NO:2)

which is 44% and 51% identical to the corresponding regions of fly and butterfly *ptc* respectively.

Using an alignment of the four insect homologues in the first hydrophilic loop of the *ptc*, two PCR primers were designed to a five and six amino acid stretch which were identical and of low degeneracy. These primers were used to isolate the mouse homologue using RT-PCR on embryonic limb bud RNA. An appropriately sized band was amplified and upon cloning and sequencing, it was found to encode a protein 65% identical to fly *ptc*. Using the cloned PCR product and subsequently, fragments of mouse *ptc* cDNA, a mouse embryonic cDNA library was screened. From about 300,000 plaques, 17 clones were identified and of these, 7 form overlapping cDNA's that comprise most of the protein-coding sequence (SEQ ID NO:9) .

Developmental and Tissue Distribution of Mouse ptc RNA. In both the embryonic and adult Northern blots, the *ptc* probe detects a single 8kb message. Further exposure does not reveal any additional minor bands. Developmentally, *ptc* mRNA is present in low levels as early as 7 dpc and becomes quite abundant by 11 and 15 dpc. While the gene is still present at 17 dpc, the Northern blot indicates a clear decrease in the amount of message at this stage. In the adult, *ptc* RNA is present in high amounts in the brain and lung, as well as in moderate amounts in the kidney and liver. Weak signals are detected in heart, spleen, skeletal muscle, and testes.

In situ Hybridization of Mouse ptc in Whole and Section Embryos. Northern analysis indicates that *ptc* mRNA is present at 7 dpc, while there is no detectable signal in sections from 7.75 dpc embryos. This discrepancy is explained by the low level of transcription. In contrast, *ptc* is present at high levels along the neural axis of 8.5 dpc embryos. By 11.5 dpc, *ptc* can be detected in the developing lung buds and gut, consistent with its adult Northern profile. In addition, the gene is present at high levels in the ventricular zone of the central nervous system, as well as in the zona limitans of the prosencephalon. *ptc* is also strongly transcribed in the condensing cartilage of 11.5 and 13.5 dpc limb buds, as well as in the ventral portion of the somites, a region which is prospective sclerotome and eventually forms bone in the vertebral column. *ptc* is present in a wide range of tissues from endodermal, mesodermal and ectodermal origin supporting its fundamental role in embryonic development.

Isolation of the Human ptc Gene. To isolate human *ptc* (h*ptc*), 2 x 105 plaques from a human lung cDNA library (HL3022a, Clonetech) were screened with a lkbp mouse *ptc* fragment, M2-2. Filters were hybridized overnight at reduced stringency (60° C in 5X SSC, 10% dextran sulfate, 5X Denhardt's, 0.2 mg/ml sonicated salmon sperm DNA, and 0.5% SDS). Two positive plaques (Hl and H2) were isolated, the inserts cloned into pBluescript, and upon sequencing, both contained sequence highly similar to the mouse *ptc* homolog. To

isolate the 5' end, an additional 6 x 105 plaques were screened in duplicate with M2-3 EcoRI and M2-3 Xho I (containing 5' untranslated sequence of mouse *ptc*) probes. Ten plaques were purified and of these, inserts were subcloned into pBluescript. To obtain the full coding sequence, H2 was fully and H14, H20, and H21 were partially sequenced. The 5.lkbp of human *ptc* sequence (SEQ ID NO:18) contains an open reading frame of 1447 amino acids (SEQ ID NO:19) that is 96% identical and 98% similar to mouse *ptc*. The 5' and 3' untranslated sequences of human *ptc* (SEQ ID NO:18) are also highly similar to mouse *ptc* (SEQ ID NO:19) suggesting conserved regulatory sequence.

Comparison of Mouse, Human, Fly and Butterfly Sequences. The deduced mouse *ptc* protein sequence (SEQ ID NO:10) has about 38% identical amino acids to fly *ptc* over about 1,200 amino acids. This amount of conservation is dispersed through much of the protein excepting the C-terminal region. The mouse protein also has a 50 amino acid insert relative to the fly protein. Based on the sequence conservation of *ptc* and the functional conservation of *hedgehog* between fly and mouse, one concludes that *ptc* functions similarly in the two organisms. A comparison of the amino acid sequences of mouse (m*ptc*) (SEQ ID NO:10), human (h*ptc*) (SEQ ID NO:19), butterfly (b*ptc*)(SEQ ID NO:4) and *drosophila* (*ptc*) (SEQ ID NO:6) is shown in the follwing table.

ALIGNMENT OF HUMAN, MOUSE, FLY, AND BUTTERFLY *PTC* HOMOLOGS

```
HPTC    MASAGNAAEPQDR--GGGGSGCIGAPGRPAGGGRRRRTGGLRRAAAPDRDYLHRPSYCDA
MPTC    MASAGNAA---------------GALGRQAGGGRRRRTGGPHRA-APDRDYLHRPSYCDA
PTC     M-----DRDSLPRVPDTHGD--VVDE---------KLFSDL---------YI-RTSWVDA
BPTC    MVAPDSEAPSNPRITAAHESPCATEA---------RHSADL---------YI-RTSWVDA
        *                                  .   ..         *. * *  **

HPTC    AFALEQISKGKATGRKAPLWLRAKFQRLLFKLGCYIQKNCGKFLVVGLLIFGAFAVGLKA
MPTC    AFALEQISKGKATGRKAPLWLRAKFQRLLFKLGCYIQKNCGKFLVVGLLIFGAFAVGLKA
PTC     QVALDQIDKGKARGSRTAIYLRSVFQSHLETLGSSVQKHAGKVLFVAILVLSTFCVGLKS
BPTC    ALALSELEKGNIEGGRTSLWIRAWLQEQLFILGCFLQGDAGKVLFVAILVLSTFCVGLKS
           **  .. **.   * .... .*. .*   *   **     *  . ** * *..*....* ****.

HPTC    ANLETNVEELWVEVGGRVSRELNYTRQKIGEEAMFNPQLMIQTPKEEGANVLTTEALLQH
MPTC    ANLETNVEELWVEVGGRVSRELNYTRQKIGEEAMFNPQLMIQTPKEEGANVLTTEALLQH
PTC     AQIHSKVHQLWIQEGGRLEAELAYTQKTIGEDESATHQLLIQTTHDPNASVLHPQALLAH
BPTC    AQIHTRVDQLWVQEGGRLEAELKYTAQALGEADSSTHQLVIQTAKDPDVSLLHPGALLEH
        *.... *..**.. ***.   ** **  .  .**        **.***   .    ..*    *** *

HPTC    LDSALQASRVHVYMYNRQWKLEHLCYKSGELITET-GYMDQIIEYLYPCLIITPLDCFWE
MPTC    LDSALQASRVHVYMYNRQWKLEHLCYKSGELITET-GYMDQIIEYLYPCLIITPLDCFWE
PTC     LEVLVKATAVKVHLYDTEWGLRDMCNMPSTPSFEGIYYIEQILRHLIPCSIITPLDCFWE
BPTC    LKVVHAATRVTVHMYDIEWRLKDLCYSPSIPDFEGYHHIESIIDNVIPCAIITPLDCFWE
        *      *. * * .*. .* * ..*   ..     *     .. *. . ** ***********

HPTC    GAKLQSGTAYLLGKPPLR----WTNFDPLEFLEELK------KINYQVDSWEEMLNKAEV
MPTC    GAKLQSGTAYLLGKPPLR----WTNFDPLEFLEELK------KINYQVDSWEEMLNKAEV
PTC     GSQLL-GPESAVVIPGLNQRLLWTTLNPASVMQYMKQKMSEEKISFDFETVEQYMKRAAI
BPTC    GSKLL-GPDYPIYVPHLKHKLQWTHLNPLEVVEEVK-KL---KFQFPLSTIEAYMKRAGI
          *..* *      *  * *       ** ..*    .. .*       *.  .    . *  ...*.
```

```
HPTC   GHGYMDRPCLNPADPDCPATAPNKNSTKPLDMALVLNGGCHGLSRKYMHWQEELIVGGTV
MPTC   GHGYMDRPCLNPADPDCPATAPNKNSTKPLDVALVLNGGCQGLSRKYMHWQEELIVGGTV
PTC    GSGYMEKPCLNPLNPNCPDTAPNKNSTQPPDVGAILSGGCYGYAAKHMHWPEELIVGGRK
BPTC   TSAYMKKPCLDPTDPHCPATAPNKKSGHIPDVAAELSHGCYGFAAAYMHWPEQLIVGGAT
        .** .***.* .*.** *****.*     *..  *. ** *  .   *** *.*****

HPTC   KNSTGKLVSAHALQTMFQLMTPKQMYEHFKGYEYVSHINWNEDKAAAILEAWQRTYVEVV
MPTC   KNATGKLVSAHALQTMFQLMTPKQMYEHFRGYDYVSHINWNEDRAAAILEAWQRTYVEVV
PTC    RNRSGHLRKAQALQSVVQLMTEKEMYDQWQDNYKVHHLGWTQEKAAEVLNAWQRNFSREV
BFTC   RNSTSALRSARALQTVVQLMGEREMYEYWADHYKVHQIGWNQEKAAAVLDAWQRKFAAEV
       .* .. *  *.***.. ***  ..**.   .    * .. * ...** .*.****  .  *

HPTC   HQSVAQNSTQK----VLSFTTTTLDDILKSFSDVSVIRVASGYLLMLAYACLTMLRW-DC
MPTC   HQSVAPNSTQK----VLPFTTTTLDDILKSFSDVSVIRVASGYLLMLAYACLTMLRW-DC
PTC    EQLLRKQSRIATNYDIYVFSSAALDDILAKFSHPSALSIVIGVAVTVLYAFCTLLRWRDP
BPTC   RKI-TTSGSVSSAYSFYPFSTSTLNDILGKFSEVSLKNIILGYMFMLIYVAVTLIQWRDP
        .      .            *....*.***  **. *    .  *     . *   *...* *

HPTC   SKSQGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLALGVGVDDVFLLAHAF
MPTC   SKSQGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLALGVGVDDVFLLAHAF
PTC    VRGQQSSVGVAGVLLMCFSTAAGLGLSALLGIVFNAASTQVVPFLALGLGVDHIFMLTAAY
BPTC   IRSQAGVGIAGVLLLSITVAAGLGFCALLGIPFNASSTQIVPFLALGLGVQDMFLLTHTY

HPTC   SETGQNKRIPFEDRTGECLKRTGASVALTSISNVTAFFMAALIPIPALRAFSLQAAVVVV
MPTC   SETGQNKRIPFEDRTGECLKRTGASVALTSISNVTAFFMAALIPIPALRAFSLQAAVVVV
PTC    AESN------RREQTKLILKKVGPSILFSACSTAGSFFAAAFIPVPALKVFCLQAAIVMC
BPTC   VEQAGD--VPREERTGLVLKKSGLSVLLASLCNVMAFLAAALLPIPAFRVFCLQAAILLL

HPTC   FNFAMVLLIFPAILSMDLYRREDRRLDIFCCFTSPCVSRVIQVEPQAYTDTHDNTRYSPP
MPTC   FNFAMVLLIFPAILSMDLYRPEDRRLDIFCCFTSPCVSRVIQVEPQAYTEPHSNTRYSPP
PTC    SNLAAALLVFPAMISLDLRRRTAGRADIFCCCF-PVWKEQPKVAPPVLPLNNNNGR----
BPTC   FNLGSILLVFPAMISLDLRRRSAAPADLLCCLM-P---ESP------LPKKKIPER----

HPTC   PPYSSHSFAHETQITMQSTVQLRTEYDPHTHVYYTTAEPRSEISVQPVTVTQDT LSCQSP
MPTC   PPYTSHSFAHETHITMQSTVQLRTEYDPHTHVYYTTAEPRSEISVQPVTVTQDNLSCQSP
PTC    ------ -------------------------- GARHPKSCNNNRVPLPAQNPLLEQPA
BPTC   --------------------------------- AKTRKNDKTHRID-TTRQPLDPDVS

HPTC   ESTSSTRDLLSQFSDSSLHCLEPPCTKWTLSSFAEKHYAPFLLKPKAKVVVIFLFLGLLG
MPTC   ESTSSTRDLLSQFSDSSLHCLEPPCTKWTLSSFAEKHYAPFLLKPKAKVVVILLFLGLLG
PTC    DIPGSS------------HSLASF----SLATFAFQHYTPFLMRSWVKFLTVMGFLAALI
BPTC   ENVTKT------------CCL-SV----SLTKWAKNQYAPFIMRPAVKVTSMLALIAVIL

HPTC   VSLYGTTRVRDGLDLTDIVPRETREYDFIAAQFKYFSFYNMYIVTQKA-DYPNIQHLLYD
MPTC   VSLYGTTRVRDGLDLTDIVPRETREYDFIAAQFKYFSFYNMYIVTQKA-DYPNIQHLLYD
PTC    SSLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTRLFGFYSMYAVTQGNFEYPTQQQLLRD
BPTC   TSVWGATKVKDGLDLTDIVPENTDEHEFLSRQEKYFGFYNMYAVTQGNFEYPTNQKLLYE

HPTC   LHRSFSNVKYVMLEENKQLPKMWLHYFRDWLQGLQDAFDSDWETGKIMPNN-YKNGSDDG
MPTC   LHKSFSNVKYVMLEENKQLPQMWLHYFRDWLQGLQDAFDSDWETGRIMPNN-YKNGSDDG
PTC    YHDSFVRVPHVIKNDNGGLPDFWLLLFSEWLGNLQKIFDEEYRDGRLTKECWFPNASSDA
BPTC   YHDQFVRIPNIIKNDNGGLTKFWLSLFRDWLLDLQVAFDKEVASGCITQEYWCKNASDEG
```

```
HPTC   VLAYKLLVQTGSRDKPIDISQLTK-QRLVDADGIINPSAFYIYLTAWVSNDPVAYAASQA
MPTC   VLAYKLLVQTGSRDKPIDISQLTK-QRLVDADGIINPSAFYIYLTAWVSNDPVAYAASQA
PTC    ILAYKLIVQTGHVDNPVDKELVLT-NRLVNSDGIINQRAFYNYLSAWATNDVFAYGASQG
BPTC   ILAYKLMVQTGHVDNPIDKSLITAGHRLVDKDGIINPKAFYNYLSAWATNDALAYGASQG

HPTC   NIRPHRPEWVHDKADYMPETRLRIPAAEPIEYAQFPFYLNGLRDTSDFVEAIEKVRTICS
MPTC   NIRPHRPEWVHDKADYMPETRLRIPAAEPIEYAQFPFYLNGLRDTSDFVEAIEKVRVICN
PTC    KLYPEPRQYFHQPNEY----DLKIPKSLPLVYAQMPFYLHGLTDTSQIKTLIGHIRDLSV
BPTC   NLKPQPQRWIHSPEDV----HLEIKKSSPLIYTQLPFYLSGLSDTDSIKTLIRSVRDLCL

HPTC   NYTSLGLSSYPNGYPFLFWEQYIGLPHWLLLFISVVLACTFLVCAVFLLNPWTAGIIVMV
MPTC   NYTSLGLSSYPNGYPFLFWEQYISLRHWLLLSISVVLACTFLVCAVFLLNPWTAGIIVMV
PTC    KYEGFGLPNYPSGIPFIFWEQYMTLRSSLAMILACVLLAALVLVSLLLLSVWAAVLVILS
BPTC   KYEAKGLPNFPSGIPFLFWEQYLYLRTSLLLALACALGAVFIAVMVLLLNAWAAVLVTLA

HPTC   LALMTVELFGMMGLIGIKLSAVPVVILIASVGIGVEFTVHVALAFLTAIGDKNRRAVLAL
MPTC   LALMTVELFGMMGLIGIKLSAVPVVILIASVGIGVEFTVHVALAFLTAIGDKNHRAMLAL
PTC    VLASLAQIFGAMTLLGIKLSAIPAVILILSVGMMLCFNVLISLGFMTSVGNRQRRVQLSM
BPTC   LATLVLQLLGVMALLGVKLSAMPPVLLVLAIGRGVHFTVHLCLGFVTSIGCKRRRASLAL

HPTC   EHMFAPVLDGAVSTLLGVLMLAGSEFDFIVRYFFAVLAILTILGVLNGLVLLPVLLSFFG
MPTC   EHMFAPVLDGAVSTLLGVLMLAGSEFDFIVRYFFAVLAILTVLGVLNGLVLLPVLLSFFG
PTC    QMSLGPLVHGMLTSGVAVFMLSTSPFEFVIPHFCWLLLVVLCVGACNSLLVFPILLSMVG
BPTC   ESVLAPVVHGALAAALAASMLA.ASEFGFVARLFLRLLLALVFLGLIDGLLFFPIVLSILO

HPTC   PYPEVSPANGLNRLPTPSPEPPPSVVRFAMPPGHTHSGSDSSDSEYSSQTTVSGLSE-EL
MPTC   PCPEVSPANGLNRLPTPSPEPPPSVVRFAVPPGHTNNGSDSSDSEYSSQTTVSGISE-EL
PTC    PEAELVPLEHPDRISTPSPLPVRSSKRSGKSYVVQGSRSSRGSCQKSHHHHHKDLNDPSL
BPTC   PAAEVRPIEHPERLSTPSPKCSPIHPRKSSSSSGGGDKSSRTS--KSAPRPC----APSL

HPTC   RHYEAQQGAGGPAHQVIVEATENPVFAHSTVVHPESRHHPPSNPRQQPHLDSGSLPPGRQ
MPTC   RQYEAQQGAGGPAHQVIVEATENPVFARSTVVHPDSPHQPPLTPRQQPHLDSGSLSPGRQ
PTC    TTITEEPQSWKSSNSSIQMPNDWTYQPREQ--RPASYAAPPPAYHKAAAQQHHQHQGPPT
BPTC   TTITEEPSSWHSSAHSVQSSMQSIVVQPEVVVETTTYNGSDSASGRSTPTKSSHGGAITT

HPTC   GQQPRRDPPREGLWPPLYRPRRDAFEISTEGHSGPSNRARWGPRGARSHNPPNPASTAMG
MPTC   GQQPRRDPPREGLRPPPYRPRRDAFEISTEGHSGPSNRDRSGPRGARSHNPRNPTSTAMG
PTC    TPPPPFPTA-----------------YPPELQSIVVQPEVTVETTHS----------DS
BPTC   TKVTATANIKVEVVTPSDRKSRRSYHYYDRRRDRDEDRDRDRERDRDRDRDRDRDRDRDR

HPTC   SSVPGYCQPITTVTASASVTVAVHPPPVPGPGRNPRGGLCPGY---PETDHGLFEDPHVP
MPTC   SSVPSYCQPITTVTASASVTVAVHPP--PGPGRNPRGGPCPGYESYPETDHGVFEDPHVP
PTC    NT--------TKVTATANIKVELAMP-----GPAVRS---YNFTS--------------
BPTC   DR--------DRERSRERDRP.DRYRD-----EPDHPA---SPRENGRDSGHE-------

HPTC   FHVRCERRDSKVEVIELQDVECEERPRGSSSN
MPTC   FHVRCERRDSKVEVIELQDVECEERPWGSSSN
PTC    --------------------------------
BPTC   -------------------------- SDSSRH
```

The identity of ten other clones recovered from the mouse library is not determined. These cDNAs cross-hybridize with mouse *ptc* sequence, while differing as to their restriction maps. These genes encode a family of proteins related to the patched protein. Alignment of the human and mouse nucleotide sequences, which includes coding and noncoding sequence, reveals 89% identity.

Radiation hybrid mapping of the human ptc gene. Oligonucleotide primers and conditions for specifically amplifying a portion of the human *ptc* gene from genomic DNA by the polymerase chain reaction were developed. This marker was designated STS SHGC-8725. It generates an amplification product of 196 bp, which is observed by agarose gel electrophoresis when o human DNA is used as a template, but not when rodent DNA is used. Samples were scored in duplicate for the presence or absence of the 196 bp product in 83 radiation hybrid DNA samples from the Stanford G3 Radiation Hybrid Panel (purchased from Research Genetics, Inc.) By comparison of the pattern of G3 panel scores for those with a series of Genethon meiotic linkage 5 markers, it was determined that the human *ptc* gene had a two point lod score of 1,000 with the meiotic marker D9S287, based on no radiation breaks being observed between the gene and the marker in 83 hybrid cell lines. These results indicate that the *ptc* gene lies within 50-100 kb of the marker. Subsequent physical mapping in YAC and BAC clones confirmed this close linkage estimate. Detailed map information can be obtained from http://www.shgc.stanford.edu.

Analysis of BCNS mutations. The basal cell nevus syndrome has been mapped to the same region of chromosome 9q as was found for *ptc*. An initial screen of EcoRl digested DNA from probands of 84 BCNS kindreds did not reveal major rearrangements of the *ptc* gene, and so screening was performed for more subtle sequence abnormalities. Using vectorette PCR, by the method according to Riley *et al.* (1990) N.A.R. 18:2887-2890, on a BAC that contains genomic DNA for the entire coding region of *ptc*, the intronic sequence flanking 20 of the 24 exons was determined. Single strand conformational polymorphism analysis of PCR-amplified DNA from normal individuals, BCNS o patients and sporadic basal cell carcinomas (BCC) was performed for 20 exons of *ptc* coding sequence. The amplified samples giving abnormal bands on SSCP were then sequenced.

In blood cell DNA from BCNS individuals, four independent sequence changes were found; two in exon 15 and two in exon 1 0. One 49 year old man was found to have a sequence change in exon 15. His affected sister and daughter have the same alteration, but three unafflicted relatives do not. His blood cell DNA has an insertion of 9 base pairs at nucleotide 2445 of the coding sequence, resulting in the insertion of three amino acids (PNI) after amino acid 815. Because the normal sequence preceding the insertion is also PNI, a direct repeat has been formed.

The second case of an exon 15 change is an 18 year old woman who developed jaw cysts at age 9 and BCCs at age 6. The developmental effects together with the BCCs indicate that she has BCNS, although none of her relatives are known to have the syndrome. Her blood cell DNA has a deletion of 11 bp, removing the sequence ATATCCAGCAC at nucleotides 2441 to 2452 of the coding sequence. In addition, nucleotide 2452 is changed from a T to an A. The deletion results in a frameshift that is predicted to truncate the protein after amino acid 813 with the addition of 9 amino acids. The predicted mutant protein is truncated after the seventh transmembrane domain. In *Drosophila, a ptc* protein that is truncated after the sixth transmembrane domain is inactive when ectopically expressed, in contrast to the full-length protein, suggesting that the human protein is inactivated by the exon 15 sequence change. The patient with this mutation is the first affected family member, since her parents, age 48 and 50, have neither BCCs nor other signs of the BCNS- DNA from both parents' genes have the normal nucleotide sequence for exon 15, indicating that the alteration in exon 15 arose in the same generation as did the BCNS phenotype. Hence her disease is the result of a new mutation. This sequence change is not detected in 84 control chromosomes.

Analysis of sporadic basal cell carcinomas. To determine whether *ptc* is also involved in BCCs that are not associated with the BCNS or germline changes, DNA was examined from 12 sporadic BCCS. Three alterations were found in these tumors. In one tumor, a C to T transition in exon 3 at nucleotide 523 of the coding sequence changes a highly conserved leucine to phenylalanine at residue 175 in the first putative extracellular loop domain Blood cell DNA from the same individual does not have the alteration, suggesting that it arose somatically in the tumor. SSCP was used to examine exon 3 DNA from 60 individuals who do not have BCNS, and found no changes from the normal sequence. Two other sporadic BCCs have deletions o encompassing exon 9 but not extending to exon 8.

The existence of sporadic and hereditary forms of BCCs is reminiscent of the characteristics of the two forms of retinoblastoma. This parallel, and the frequent deletion in tumors of the copy of chromosome 9q predicted by linkage to carry the wild-type allele, demonstrates that the human *ptc* is a tumor suppressor gene. *ptc* represses a variety of genes, including growth factors, during Drosophila development and may have the same effect in human skin. The often reported large body size of BCNS patients also could be due to reduced *ptc* function, perhaps due to loss of control of growth factors. The C to T transition identified in *ptc* in the sporadic BCC is also a common genetic change in the *p53* gene in BCC and is consistent with the role of sunlight in causing these tumors. By contrast, the inherited deletion and insertion mutations identified in BCNS patients, as expected, are not those characteristic of ultraviolet mutagenesis.

The identification of the *ptc* mutations as a cause of BCNS links a large body of developmental genetic information to this important human disease. In embryos lacking *ptc* function part of each body segment is transformed into an anterior-posterior mirror-image duplication of another part. The patterning changes in *ptc* mutants are due in part to derepression of another segment polarity gene, *wingless,* a homolog of the vertebrate Wnt genes that encodes secreted signaling proteins. In normal embryonic development, *ptc* repression of *wg* is relieved by the Hh signaling protein, which emanates from adjacent cells in the posterior part of each segment. The resulting localized wg expression in each segment primordium organizes the pattern of bristles on the surface of the animal. The *ptc* gene inactivates its own transcription, while Hh signaling induces *ptc* transcription.

In flies two other proteins work together with Hh to activate target genes: the ser/thr kinase *fused* and the zinc finger protein encoded by *cubitus interruptus.* Negative regulators working together with *ptc* to repress targets are *protein kinase A and costal2.* Thus, mutations that inactivate human versions of *protein kinase A or costal2,* or that cause excessive activity of human *hh, gli,* or a *fused* homolog, may modify the BCNS phenotype and be important in tumorigenesis.

In accordance with the subject invention, mammalian patched genes, including the mouse and human genes, are provided, which can serve many purposes. Mutations in the gene are found in patients with basal cell nevus syndrome, and in sporadic basal cell carcinomas. The autosomal dominant inheritance of BCNS indicates that *patched* is a tumor suppressor gene. The patched protein may be used in a screening for agonists and antagonists, and for assaying for the transcription of *ptc* mRNA. The protein or fragments thereof may be used to produce antibodies specific for the protein or specific epitopes of the protein. In addition, the gene may be employed for investigating embryonic development, by screening fetal tissue, preparing transgenic animals to serve as models, and the like.

As described above, patients with basal cell nevus syndrome have a high incidence of multiple basal cell carcinomas, medulloblastomas, and meningiomas. Because somatic *ptc* mutations have been found in sporadic basal cell carcinomas, we have screened for *ptc* mutations in several types of sporadic extracutaneous tumors. We found that 2 of 14 sporadic medulloblastomas bear somatic nonsense mutations in one copy of the gene and also deletion of the other copy. In addition, we identified mis-sense mutations in *ptc* in two of seven breast carcinomas, one of nine meningiomas, and one colon cancer cell line. No *ptc* gene mutations were detected in 10 primary colon carcinomas and eighteen bladder carcinomas.

BCNS[3] (OMIM #109400) is a rare autosomal dominant disease with diverse phenotypic abnormalities, both tumorous (BCCs, medulloblastomas, and meningiomas) and developmental (misshapen ribs, spina bifida occults, and skull abnormalities; Gorlin, R.J. (1987) Medicine 66:98-113). The BCNS gene was mapped to chromosome 9q22.3 by

linkage analysis of BCNS families and by LOH analysis in sporadic BCCs (Gallani, M.R. *et al.* (1992) Cell 69:111-117). LOH in sporadic medulloblastomas has been reported in the same chromosome region (Schofield, D. *et al.* (1995) Am J Pathol 146:472-480). Recently, the human homologue of the *Drosophila patched* (PTCII) gene has been mapped to the BCNS region (Hahn, H. *et al.* (1996) Cell 85:841-851; Johnson, R.L. *et al.* (1996) Science 272:1668-1671; Gallani, M.R. *et al.* (1996) Nat Genet 14:78-81; Xie, J. *et al.* (1997) Genes Chromosomes Cancer 18:305-309), and mutations in this gene have been found in the blood DNA of BCNS patients and in the DNA of sporadic BCCs (Hahn, H. *et al., supra;* Johnson, R.L. *et al., supra*; Gallani, M.R. *et al., supra*; and Chidambaram, A. *et al.* (1996) Cancer Res 36:4599-4601). *ptc* appears to function as a tumor suppressor gene; inactivation abrogates its normal inhibition of the hedgehog signaling pathway. Because of the wide variety of tumors in patents with the BCNS and wide tissue distribution of *ptc* gene expression, we have begun screening for *ptc* gene mutations in several types of human cancers, especially those present in increased numbers in BCNS patients (medulloblastomas), those in tissues derived embryologically from epidermis (breast carcinomas) and those with chromosome 9q LOG (bladder carcinomas; see Cairns, P. *et al.* (1993) Cancer Res 53:1230-1232; and Sidransky, D. *et al.* (1997) NEJM 326:737-740).

To further study the roles of *ptc* in development and in tumorigenesis, we have constructed mice lacking *ptc* function. By homologous recombination, part of *ptc* exon 1 (including the putative start codon) and all of exon 2 were replaced with *lacZ* and a neomycin resistance gene (Fig. 3) (DNA from the *ptc* genomic locus was isolated from a 129SV genomic phage library [Stratagene]. Exons 1-15 of human *PTC (1)* were mapped by PCR and sequencing. The 3' arm of homology was a 3.5 kb EcoRI-BamHI fragment from the second intron that gained a BamHI site from pBSII [Stratagene] and was cloned into the BamHI site of pPNT [Tybulewicz, *et al.* (1991) Cell 65:1153]. A cassette containing the gene for nuclear localized b-galactosidase, followed by the mP1 intron and polyA tail was excised from pNLacF [Mercer, *et al.* (1991) Neuron 7:703] and cloned into the Xho I site of pPNT using Xho I and Sal I linkers. The 5' arm of homology was a 6.5 kb Xho I to Nru I fragment that was cloned into the Xho I site upstream of *lacZ* via a Sal I linker. The Nru I site is in the first *ptc* exon. The resulting plasmid, KO1, was linearized with Xho I and electroporated into RI ES cells that were subjected to double selection and analyzed by Southern blot [Joyner, A.L. Gene Targeting: A Practical Approach. Oxford University Press, New York, 1993, pp.33-61]. Targeted clones were expanded and used for injection into C57Bl/6 blastocysts [Hogan, B. *et al.* Manipulating the Mouse Embryo: A Laboratory Manual Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1994, pp.196-204]. Protein made from any alternative ATG would lack the first proposed transmembrane domain, flipping the orientation of the protein in the membrane. Three independent ES clones were used to make chimeras that were bred to B6D2F1 animals to generate

heterozygous mice on a mixed background. Interbreeding of heterozygotes produced no homozygous animals among 202 offspring examined. Analysis of embryos from timed matings suggested that $ptc^{-/-}$ embryos die between embryonic day (E) 9.0 and E10.5, with the first gross phenotypes appearing by E8. In $ptc^{-/-}$ embryos, the neural tube failed to close completely and was overgrown in the head folds, hindbrain and spinal cord (Fig. 4, A to C). Embryonic lethality may have been due to abnormal development of the heart (Fig. 4B), which never beats.

In flies Ptc protein inhibits *ptc* transcription. By inhibiting Ptc function, Hh increases production of Ptc which may then bind available Hh and limit the range or duration of effective Hh signal (Y. Chen and G. Struhl, (1996) Cell 87:553). Hh signaling also post-transcriptionally regulates the zinc finger protein cubitus interruptus (ci) (C. K. Motzny and R. Holmgren, (1996) Mech Dev 52:137; Domínguez, *et al.* (1996) Science 272:1621; Hepker, *et al.* (1997) Development 124:549; Aza-Blanc, *et al.*, (1997) Cell 89:1043). In vertebrates, Sonic hedgehog (Shh) signaling induces transcription of both *ptc* and a *ci* homolog, *Gli* (Goodrich, *et al.* (1996) Genes Devel. 10:301; Marigo, *et al.* (1996) Development 122:1225; Concordet, *et al.*, (1996) Development 122:2835; Marigo, *et al.* (1996) Dev. Biol. 180:273). Derepression of *ptc* and *Gli* in $ptc^{-/-}$ mice should therefore reveal where Ptc is normally active.

ptc and *Gli* expression was greatly increased in $ptc^{-/-}$ embryos. In $ptc^{+/-}$ mice expression of the *lacZ* gene fused to the first *ptc* exon during targeting accurately reported the pattern of *ptc* transcription (Fig. 4, C and D). In $ptc^{-/-}$ embryos expression of *ptc-lacZ* was extensively derepressed starting at about E8.0 in the anterior neural tube and spreading posteriorly by E8.75 (Fig. 4, C and E). Derepression was germ layer-specific: both *ptc-lacZ* and *Gli* were expressed throughout the ectoderm and mesoderm, but not in the endoderm (Fig. 4, D to G). *ptc* expression may be excluded from the endoderm in order to avoid interfering with Shh signaling from the endoderm to the mesoderm (Roberts *et al.*, (1995) Development 121:3163). A differential requirement for Ptc may distinguish the germ layers.

As revealed by *ptc* mutants, an early site of Ptc activity is the neural tube, where Shh and Ptc act antagonistically to determine cell fates. Shh induces the floor plate and motor neurons in the ventral neural tube (Echelard *et al.*, (1993) Cell 75:1417; Roelink *et al.*, (1994) Cell 76:761; Roelink *et al.*, (1995) Cell 81:445-455). These cell types fail to form in *Shh* mutants (Chiang *et al.*, (1996) Nature 383:407). High levels of Shh produced by the notochord may induce floor plate by completely inactivating Ptc (Echelard *et al.*, (1993) *supra*; Roelink *et al.*, (1994) *supra*; Roelink *et al.*, (1995) *supra*). If so, elimination of *ptc* function might cause floor plate differentiation throughout the neural tube. Prospective floor plate cells transcribe the forkhead transcription factor *HNF3b* first and then *Shh* itself (Echelard *et al.*, (1993) *supra*; Roelink *et al.*, (1994) *supra*; Roelink *et al.*, (1995) *supra*). In

E8.5 *ptc* mutants, transcription of *HNF3b* and *Shh* was expanded dorsally (Fig. 5, A to C). Ectopic *Shh* expression was most extensive in the anterior, where transcripts could be detected throughout the neurepithelium (Fig. 5, B and C). Cells in this region were in a single layer with basal nuclei, like floor plate cells that are normally restricted to the ventral midline (Fig. 5, D and E). Expression of the intermediate neural tube marker *Pax6* (C. Walther and P. Gruss, (1991) Development 113:1435) was completely absent from *ptc* mutant embryos, suggesting that only ventral, and not ventrolateral, cell fates are specified (Fig. 5, F and G).

Dorsalizing signals from the surface ectoderm (Dickinson, *et al.* (1995) Development 121:2099; Liem, *et al.* (1995) Cell 82:969) could confer dorsal cell fates even in the absence of *ptc* function. In E8-E9 *ptc* homozygotes the dorsal neural tube marker *Pax3* was not expressed in the anterior neural tube, but was transcribed in a very small region at the dorsal-most edge of the posterior neural tube (Fig. 5, H to J). In addition *erb-b3* transcription, which marks migratory neural crest cells (Fig. 5K) (H. U. Wang and D. J. Anderson, (1997) Neuron 18:383), was not detected in the somites of *ptc* mutants (Fig. 5L). We conclude that only limited dorsal fate determination occurs in the absence of *ptc*. BMP signals maintain dorsal gene expression (Dickinson, *et al.* (1995) *supra*; Liem, *et al.* (1995) *supra*), so either *ptc* is required for BMPs to work or BMP signaling is ineffective in most cells expressing Shh targets.

Ventralization of the neural tube in *ptc* mutants occurred without affecting cell identity along the rostrocaudal axis. In *ptc*$^{-/-}$ embryos, cells in the anterior neural tube expressed the forebrain marker *Nkx2.1* (Shimamura, *et al.* (1995) Development 121:3923) and cells in the spinal cord transcribed low levels of *hoxb1* (Wilkinson, *et al.* (1989) Nature 341:405) (Fig. 5, M and N). *hoxb1* was not transcribed in the fourth rhombomere of *ptc* mutants (Fig. 5, N). This may reflect a transformation of hindbrain cells to floor plate, since *hoxb1* is excluded from the midline of wild-type embryos. Conversely, in the anterior, *Nkx2.1* expression was expanded dorsally in mutants compared to wild-type embryos (Fig. 5, M).

ptc$^{+/-}$ mice had phenotypes similar to those of BCNS patients: they were larger than their wild-type littermates [30.72 $\pm$ 3.83 (average $\pm$ SD; n=29) vs. 26.54 $\pm$ 2.51 (n=39) at 2-3 months; *P*=0.000001], a small fraction (3 of 389 mice examined) had hindlimb defects such as extra digits or syndactyly (Fig. 6A) or obvious soft tissue tumors (1 of 243) and many developed brain tumors (see below).

Of 243 *ptc*$^{+/-}$ mice which were between the ages of 2 and 9 months and were not sacrificed for other studies, 18 died or were euthanized because of sickness. No wild-type littermates died. Ten of the affected heterozygotes were autopsied and eight were found to have large growths in the cerebellum that resembled medulloblastomas (Fig. 6, B and C).

Human medulloblastomas are believed to arise from a "primitive neurectodermal" cell type (J.P. Provias and L. E. Becker, (1996) J Neurooncol 29:35). They are most common in children, can be metastatic or non-metastatic, and can have glial and neuronal properties. The histology of tumors from *ptc*$^{+/-}$mice was similar to that of human medulloblastoma: tumor cells were small, with dark carrot-shaped nuclei and little cytoplasm (Fig. 6, D and E), and although a subset expressed neurofilament protein and synaptophysin (Fig. 6F) (For immunostaining, two tumors were fixed and embedded in paraffin. Tissue sections (4 mm) were cleared and dehydrated, treated with 3% hydrogen peroxide and then with a dilution of 1:10 normal rabbit serum (Vector Laboratories). Anti-synaptophysin (Boehringer-Mannheim) was used at a dilution of 1:5 and anti-neurofilament protein (Dako) at 1:50. Antibody binding was visualized with a peroxidase Vectastain Elite ABC kit (Vector Laboratories). Nuclei were counterstained with hematoxylin. Like anti-synaptophysin, anti-neurofilament staining appeared in processes of the tumor cells.), the majority of cells appeared undifferentiated. Of the two autopsied animals without apparent medulloblastomas, one had a large tumor growing out of its rib muscle and the other died for unknown reasons. Medulloblastomas and soft tissue tumors were also observed in *ptc*$^{+/-}$ mice maintained on an inbred 129SV background: 6 of 27 had obvious medulloblastomas; 2 of 27 had tumors in the muscle of their leg; and 3 of 27 died but were not examined.

The *ptc* and *Gli* genes were strongly transcribed in the brain tumors but not in surrounding tissue (Fig. 7, A and B; n = 3 of 3 tumors examined). There was no detectable increase in *Shh* expression (Fig. 7C). To assess the incidence of medulloblastomas, brains from 47 asymptomatic *ptc*$^{+/-}$ mice were randomly collected and stained with X-gal. Nine brains contained medulloblastomas that were easily recognized by their disorganized morphology and intense *ptc-lacZ* expression (Fig. 7D). Medulloblastomas were observed in 7 of 23 (30.4%) *ptc*$^{+/-}$ mice at 12 to 25 weeks of age, 1 of 12 (8.3%) mice at 9 to 10 weeks and 1 of 12 (8.3%) mice at 5 weeks. Tumors can therefore arise as early as 5 weeks postnatally, but they increase in severity and frequency as the animal ages.

We looked for changes in *ptc-lacZ* expression that might reflect early stages of tumorigenesis. At all stages examined, about half of the animals [50% at 5 to 10 weeks (n=24), 56.5% at 12 to 25 weeks (n=23)] exhibited regions of increased X-gal staining on the surface of the cerebellum (Fig. 7E). These regions were usually lateral and often extended down into the fissures separating the folia (Fig. 7, E and F). The mouse medulloblastomas may arise from these cells, which are superficial to the molecular layer of the cerebellum (Fig. 7F). During fetal development, prospective cerebellar granule cells proliferate in the external granule layer (EGL), the outermost layer of the cerebellum. Granule cells then leave and migrate past the Purkinje cells to form the internal granule cell layer of the adult animal, gradually depleting the EGL. The remnants of the fetal EGL have been proposed to be a source of human medulloblastoma progenitors, a hypothesis consistent with the higher

frequency of these tumors in children (L. Stevenson and F. Echlin, (1934) Arch. Neurol. Psychiat. 31:93; Kadin, *et al.* (1970) J Neuropathol Exp Neurol 29:583).

The abundance of cerebellar *ptc* transcripts was reduced by about 50% in the $ptc^{+/-}$ mice compared to wild-type littermates (Fig. 7G). This reduction could lead to ectopic expression of Shh target genes and to uncontrolled cell proliferation. Brain tumors might arise from Ptc haploinsufficiency alone, from additional mutations in the second *ptc* allele, or from a combination of *ptc* mutations with mutations in other tumor suppressor loci. We have not observed BCCs in $ptc^{+/-}$ mice, perhaps because somatic inactivation of the second *ptc* gene is required as it is in human BCCs.

Our analysis has revealed that Ptc controls growth and pattern formation in early neural development and in the adult cerebellum. Autoregulation of *ptc* occurs in vertebrates as it does in flies, and the balance between Hh and Ptc activities appears critical for normal development. The importance of Ptc dosage is emphasized by the phenotype of the $ptc^{+/-}$ mice, which develop a tumor type observed in the corresponding human cancer predisposition syndrome. Medulloblastoma is a common childhood brain tumor and the prognosis remains grim. The Hh/Ptc pathway may provide new diagnostic tools and new insights into tumorigenesis that may be directed toward potential therapies.

Materials and methods

Clinical Materials. Diagnoses of all tumors were confirmed histologically. Cell lines were obtained from the America Type Culture Collection. DNA was extracted from tumors or matched normal tissue (peripheral blood leukocytes or skin) as described (Cogen, P.H. *et al.* (1990) Genomics 8:279-285; and Sambrook, J. *et al.* Molecular Cloning: A Laboratory Manual, Ed. 2, Vol. 2, pp. 9.17 - 9.19, Cold Spring Harbor, NY (1989)).

PCR and Heteroduplex Analysis. PCR amplification and heteroduplex/SSCP analysis were performed as described (Johnson, R.L. *et al., supra*; Spritz, R.A. *et al.* (1992) Am J Hum Genet 51:1058-1065). Primers used and intron/exon boundary sequences of the *ptc* gene were derived as reported previously (Johnson, R.L. *et al., supra*) and are shown in Table 1. Primers for exon 1 and 2 were from Hahn *et al.* (*supra*).

Sequence Analysis. Exon segments exhibiting bands were reamplified and were sequenced directly using the Sequenase sequencing kit according to the protocol recommended by the manufacturer (United States Biochemical Corp.). A second sequencing was performed using independently amplified PCR products to confirm the sequence change. The amplified PCR products from each tumor were also cloned into the plasmid vector pCR 2.1 (InVitrogen), followed by sequence analysis of at least four independent clones. The sequence alteration was confirmed from at least two independent clones. Simplified

amplification of specific allele analysis was performed according to Lei and Hall (Lei, X. and Hall, B.G. (1994) Biotechniques 16:44-45).

Allele Loss Analysis. Microsatellites used for allelic loss analysis were D9S109, DpS119, D9S127, D9S196, and D9S287 described in the CHLC human screening set (Research Genetics). A part of the *ptc* intron 1 sequence was tested for polymorphism in a control population and found to be polymorphic in 80% of the samples tested. This microsatellite was used for analysis of *ptc* gene allelic loss in bladder carcinomas. The primer sequences are as follows: forward primer, 5'-CTGAGCAGATTTCCCAGGTC-3'; and reverse primer, 5'-CCTCAGACAGACCTTTCCTC-3'. The PCR cycling for this newly isolated marker was 4 min. at 95 C, followed by 30 cycles of 40 s at 95 C, 2 min. at 60 C, and 1 min. at 72 C. PCR products were separated on 6% polyacrylamide gels and exposed to film.

Results and Discussion

Intronic boundaries were determined for 22 exons of *ptc* by sequencing vectorette PCR products derived from BAC 192J22 (Johnson R.L., *supra*; Table 1). Our findings are in agreement with those of Hahn *et al.* (*supra*), expect that we find exon 12 is composed of 2 separate exons of 126 and 119 nucleotides. This indicates that *ptc* is composed of 23 coding exons instead of 22. In addition, we find that exons 3, 4, 10, 11, 17, 21, and 23 differ slightly in size than reported previously (Hahn *et al., supra*). Of 63 tumors studied, 14 were sporadic medulloblastomas, and 9 were sporadic meningiomas. These 23 tumors were examined for allelic deletions by genotyping of tumor and blood DNA with microsatellite markers that flank the *ptc* gene: D9S119, D9S196, D9S287, D9S127, and D9S109. Four of 14 medulloblastomas had LOH. Two of the medulloblastomas, both of which had LOH, had mutations (med34 and med36; see Cogen, P.H. *et al., supra*), which are predicted to result in truncated proteins (Table 2). DNA samples from the blood of these patients lack these mutations, indicating that they both are somatic mutations. med34 also has allelic loss on 17p (Cogen, P.H. *et al., supra*). We were unable to detect *ptc* gene mutations by heteroduplex analysis in the other two medulloblastomas bearing LOH on 9q. The pathological features of these two tumors differed in that med34 belongs to the desmoplastic subtype, whereas med36 is of the classic type, indicating that *ptc* mutations in medulloblastomas are not restricted to a specific subtype.

TABLE 1 *Primers and boundary sequences of PTCH*

Exon	5' Boundary[a]	Nucleotide position[b]	Exon size	3' boundary[a]	Reading frame[c]	Primers
1	ND[d]	ND	ND	aag|GTGNAT	ND	
2	ND	202	193	aag|GTAAGA	3	
3	TGTCAG|t	395	190	cag|GTAAGG	1	3F GAGTTTGCAGTGATGTTGCTNTTC 3R ACCGCCTTACCTGCTGCTC
4	TATTAG|a	585	70	cag|GTATAT	2	4F TGCACTAATTTTCTTATTACACTC 4R TAAGGCACACTACTGGGGTG
5	TGACAG|t	655	92	cct|GTAAGT	3	5F GAACACCCCAGTAGTGTGCC 5R TGAGTCCTAGAGAAGTCACAGACA
6	TTGCAC|a	747	199	aaa|GTGAGT	2	6F GGCTCTTTTCATGGTCTCGTC 6R TGTTTTGCTCTCCACCGTTC
7	TTTTAG|c	945	122	cag|GTAAGC	3	7F GCACTGGATTTTAACAAGGCATG 7R AGGGCATAGATTGTCCTCGG
8	CTGCAG|c	1068	148	gag|GTAAAC	2	8F TGGGAATACTGATGATGTGCC 8R CATAACCAGCGAGTCTGCAC
9	CCACAG|g	1216	132	atg|GTAACG	3	9F CATTTGGCATTTCGCATTC 9R ACCAAACCAAACTCCAGCCC
10	TTGCAG|c	1348	156	cag|GTACTA	3	10F TGCCCCCATTGTTCTGCTTG 10R GGACAGCAGATAAATGGCTCC
11	CTGTAG|g	1504	99	gag|GTAATG	3	11F GCATCTCGCATGTCTAATGCCAC 11R AAGCTGTGATGTCCCCAAAG
12	TCCCAG|g	1603	126	cag|GTGAGC	3	12F GACCATGTCCACTGCAGCTC 12R CGTTCAGGATCACCACAGCC
13	TCCCAG|g	1729	119	aag|GTACAT	3	13F AGTCCTCTGATTGGGCGGAG 13R CCATTCTGCACCCAATCAAAAC
14	TTTCAG|c	1848	403	aag|GTAATC	2	14F AAAATGGCAGAATGAAAGCACC 14R CTGATGAACTCCAAAGCTTCTG
15	TTCCAG|g	2251	310	agg|GTAAGA	3	15F GGAAGAGTCAGTGGTGCTCC 15R CGCCAAAGACCGAAAGGAC
16	TTCTAG|g	2561	143	cag|GTACTC	1	16F AGGCTCCTTCTGGCTGCGAG 16R GCTGTCAAGCAGCCTCCAC
17	TTGTAG|t	2704	184	gaa|GTAAGT	3	17F GCTCTCAAGGCAGAAGTCTG 17R GGAAGGCACCTCTCTAAGTAC
18	GTCCAG|a	2888	281	att|GTGAGT	1	18F GCTCCTAACCTGTGTCCCTTC 18R GAATTTGACTTCCACCAAAGCCC
19	CTCCAG|g	3169	138	ctg|GTATGG	3	19F CGCCCACTGACCACTGTGTG 19R GAGCCAGAGGAAATGGGTTG
20	GCACAG|g	3307	143	cag|GTAAGC	3	20F AGCATTTACCAGGTGAAGTCC 20R TTGCACACGCCTGCTTAC
21	TCCCAG|g	3450	100	gag|GTCAGT	2	21F TGTTCCCGTTTCCTCCTG 21R GCACAGGAAACACACCCATTC
22	AAATAG|g	3550	255	act|GTAAGT	3	22F GCAGGTAAATGGACAAGAACAC 22R ACTACCACGGTGGGAAGACC
23	CTGCAG|g	3805	541	gag|GTGAGT	3	23F CCCTTCTAACCCACCCTCAC 23R GACACATCAGCCTTGCTC
24	ND	4346	ND	ND		

[a] Consensus sequences for the 5' and 3' exonic boundaries are $(^{T}{}_{C})_{11}$NCAG\|g and ag\|GT$^{G}{}_{A}$AGT, respectively (20). Lower case denotes exonic sequence.
[b] Exon positions are in reference to the coding sequence of *PTCH* (3) with the beginning ATG as nucleotide 1.
[c] 5' exon boundary begins after the first, second, or third base of the codon of the translational reading frame.
[d] ND not determined.

One report (Schofield, D. *et al., supra*) has shown that five medulloblastomas (two BCNS-associated cases and three sporadic cases) bearing LOH on chromosome 9q22.3-q31 are all of the desmoplastic subtype, suggesting LOH on 9q22.3 is histological subtype specific. We feel that the conclusion derived from only five positive tumors is a not strong one because we and others (Raffel, C. *et al.* (1997) Cancer Res 57:842-845) have found nondesmoplastic subtypes of medulloblastomas bearing LOH on chromosome 9q22.3. Independently, another group has reported their finding of *ptc* mutations in sporadic medulloblastomas (Raffel, C. *et al, supra*).

A change of T to C at nucleotide 2990 (in exon 18) was identified in DNA from one of nine sporadic meningiomas, causing a predicted change of codon 997 from Ile to Thr (Table 2). The meningioma bearing this mutation also has allelic loss on 9q22.3. Blood cell DNA is heterozygous for this mutation, but DNA from the tumor contains only the mutant sequence. Of 100 normal chromosomes examined, none has this sequence change, suggesting that this mutation is not likely a common polymorphism. This patient is 84 years

old and has had no phenotypic abnormalities suggestive of the BCNS, suggesting that this sequence alteration may not have caused complete inactivation of the *ptc* gene. None of the other eight meningiomas had detectable LOH at chromosome 9q.

TABLE 2 *PATCHED gene alterations*[a]

Tumor	Pathology	Nucleotide	Codon	Exon	Consequence	LOH	Mutation Type
Med34	Medulloblastoma (desmoplastic)	TC1869A	623	14	Frameshift	Yes	Somatic
Med36	Medulloblastoma (classic)	G2503T	835	15	Glu to STOP	Yes	Somatic
Men1	Meningioma	T2990C	997	18	Ile to Thr	Yes	Germ-line
Br349	Breast carcinoma	T2863C	955	17	Tyr to His	Yes	Somatic
Br321	Breast carcinoma	A2975G	995	18	Glu to Gly	No	Somatic
Co320	Colon tumor cell line	A2000C	667	14	Glu to Ala	No	Unknown
Co8-1	Colon carcinoma	T to C	Intron 10		Polymorphism	No	Germ-line
Co15-1	Colon carcinoma	T to C	Intron 10		Polymorphism	No	Germ-line

We also examined a variety of other tumors (10 primary tumors and 1 cell line), 18 bladder tumors (14 primary tumors and 4 cell lines), and 2 ovarian cancer cell lines. These tumors are not known to occur in higher than expected frequency in BCNS patients. We identified sequence abnormalities in two breast carcinomas and in the one colon cancer cell line (Table 2). The mutation found in breast carcinoma Br349 is not present in the patient's normal skin DNA, indicating that the sequence change is a somatic mutation. Direct sequencing of the PCR product indicated that only the mutant allele is present in the tumor. This mutation changes codon 955 from Tyr to His, and this Tyr is conserved in human, murine, chicken, and fly *ptc*II homologues (Goodrich, L.V. *et al.* (1996) <u>Genes Dev</u> 10:301-312). The mutation in breast carcinoma Br321 is predicted to change codon 995 from Glu to Gly, and the tumor with this mutation retains the wild-type allele. We have sequenced exon 18 in DNA from the blood of 50 normal person s and found no changes from the published sequence, suggesting that the sequence change found in Br321 is not a common polymorphism. Furthermore, examination of the DNA from the cultured skin fibroblasts of the patient did not reveal the same mutation, indicating that this is a somatic mutation.

Because DNA is not available from normal cells of the patient from which colon cell line 320 was established, we used simplified amplification of specific allele analysis (Lei, X.

and Hall, B.G., *supra*) to examine 50 normal blood DNA samples for the presence of the sequence alteration and found none but the DNA from this cell line to have the mutant allele, suggesting that this mutation also is unlikely to be a common sequence polymorphism. For bladder carcinomas, a newly isolated microsatellite that was derived from intron 1 of the *ptc* gene was used to examine LOH in the tumor. Three primary bladder carcinomas showed LOH at this intragenic locus. With no *ptc* mutations detected in these tumors, we suspect that the LOH in these three bladder carcinomas may reflect the high incidence of while chromosome 9 loss in bladder cancers (Sidransky, D. *et al., supra*). A similar observation has been reported previously (Simoneau, A. R. *et al.* (1996) Cancer Res 56:5039-5043).

We also detected a sequence change in intron 10 in two colon carcinomas, 15-1 and 8-1, an alteration that was reported previously as a splicing mutation (Unden, A.B. *et al.* (1996) Cancer Res 56:4562-4565). Because we found the same sequence change in about 20% of normal control samples, we suggest that this more likely is a nonpathogenic polymorphism. The *ptc* protein is predicted to contain 12 transmembrane domains, two large extracellular loops, and one intracellular loop (Goodrich, L.V. *et al., supra*). Of the six mutations we identified, four are missense mutations. Three mutations lead to amino acid substitutions in the second extracellular loop, and one mutation results in an amino acid change in the intracellular domain.

Our data indicate that somatic inactivation of the *ptc* gene does occur in some sporadic medulloblastomas. In addition, because missense mutations of the *ptc* gene were detected in breast carcinomas, we suspect that defects of the *ptc* function also may be involved in some breast carcinomas, although biochemical evidence is necessary to show how these missense mutations might impair *ptc* function. Of 11 colon cancers and 18 bladder carcinomas examined, we found only one mutation in 1 colon cell line, suggesting that *ptc* gene mutations are relatively uncommon in clon and bladder cancers, although the incidence of chromosome 9 loss in bladder cancers is high (Cairns, P. *et al., supra*).

Published reports of SSCP analysis of tumor DNA identified mutations in the *ptc* gene in only 30% of sporadic BCCs, although chromosome 9q22.3 LOH was reported in more than 50% of these tumors (Gallani, M.R. *et al., supra*). It has been reported that heteroduplex/SSCP analysis of gene mutations is more sensitive than SSCP analysis (Spritz, R.A. *et al., supra*). In our studies, we were able to identify a point mutation in the 310-bp - PCR product from exon 15 using heteroduplex analysis, whereas SSCP analysis failed to reveal this sequence change (Table 2). Therefore, we suspect that there may be more mutations in BCCs than we have found thus far. Analysis of the *ptc* gene in BCNS patients and in sporadic BCCs has identified mutations scattered widely across the gene, and the majority of mutations were predicted to result in truncated proteins (Hahn, H. *et al., supra*; Johnson, R.L. *et al., supra*; Gallani, M.R. *et al., supra*; Chidambaram, A. *et al., supra*;

Unden, A.B. *et al., supra*; Wicking, C. *et al.* (1997) Am J Hum Genet 60:21-26). In our screening, we found two breast carcinomas bearing missense mutations of the *ptc* gene. In one of these two tumors, B349, direct sequencing indicated a deletion of the other copy of the *ptc* gene. Any comparison of mutations in skin cancers versus extracutaneous tumors must consider the wholly different causes of these mutations; UV light is unique to the skin.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent o application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.